

Tomorrow's conundrum of Europe

Excitement over developments in Eastern Europe should be moderated by a recognition of the chauvinistic dangers that it brings. What Europe needs is a moratorium on boundary changes. German reunification is best postponed.

MALTA's role in European affairs has not been conspicuous since the ending of the Crusades, while the meeting on (rather than, as originally intended, off) the island between Presidents Mikhail Gorbachev and George Bush will not have restored much of its influence. That it allowed the two men to smile often and openly in each other's presence is, of course, a plus; the more serious business on arms control planned for next June should now be the more easily carried through. Each, no doubt, will now also be more confident that the other is seriously seeking to divert military spending to civil causes — the US administration seems to have seized more eagerly on the chance to cut spending than other members of the North Atlantic Treaty Organization. The federal deficit plainly nags like toothache. The Soviet deficit, relatively even larger, is for the time being bridged by printing money; no wonder that Gorbachev is pleased that his new friend, Bush, may use his economic influence (but not his cheque-book) to help *perestroika* along. Whether that will help during the long winter is another matter.

What seems to have foxed the two presidents at the weekend (the weather apart) is the turmoil in Europe. They are not alone in that. Monolithic regimes in Eastern Europe have been collapsing like dominoes since the Soviet Union let it be known that it would not prevent political change in Poland, now almost a democracy if an impoverished one. But that does not mean that Gorbachev has lost control. He seems resolute in telling ancient leaders that they cannot stand in the way of history and equally determined that the East European states should stay within the Warsaw Pact, German nostalgia notwithstanding. The benefits? The expectation that the conservatives in the Soviet Union will understand that what goes for Eastern Europe also applies to them. But the pace of change has left even Gorbachev wanting for a tangible description of what he means by a "common European home". Bush is as perplexed.

It is a perplexing business. Three years from now, at the end (not the beginning) of 1992, the European Communities hope to be a true common market and thus, in due course, a dominating economic force in Europe. The chiefly French fear that the process will be interrupted by ambitions for the reunification of East and West Germany are not over-serious: the frenzy of the past few weeks will soon abate, to be replaced by mutual recognition that reunification is at best a distant prospect. Why

should two of Europe's three-and-a-third German-speaking countries (counting Switzerland ungenerously as a third) consider they have no choice but to turn the clock back almost half a century, especially when their boundaries have changed considerably since 1939 and when there are much more interesting avenues for them to explore?

The interest of Europe is its diversity. So much can be told from comparing a journey from Stockholm to Naples with one from, say, Omaha (Nebraska) to Austin (Texas), roughly the same distance. It is remarkable that the countries of Eastern Europe have kept their own distinctive character throughout the past 40 years. Now there will be even more diversity to conjure with. The danger is that its interest will be complicated by ethnic or, more accurately, linguistic tensions. Many of the new near-democracies are artificial states, glued together at Versailles and Yalta, after the First and Second World Wars. The temptation now will be to unravel these ancient packages, reuniting patches on the map of Europe as if solving a jigsaw puzzle, sometimes even carrying out gigantic programmes of 'repatriation'.

These temptations are natural, but should be suppressed by those in whom they arise. In the present heady atmosphere, people may forget that the Second World War began with hankerings of this kind. What Europe most needs now is a moratorium on external boundary changes. Internal changes are a different matter; if people thrown together in some past jigsaw game so dislike each other, fission should be allowed. Nobody would wish the present state of Belgium on, say, Czechoslovakia, but there must be better ways of letting people sense that they are in charge of their own affairs. That is why German reunification would be a bad precedent.

The better bet for everybody is that mobility, allowed for and even encouraged by the Helsinki agreements of 1979, should have a chance to make all European states more prosperous. A precondition would be that the members of the European Communities should be at least as open to other European states as they are to each other. That may be easier now that Europe's superpower hankerings have been attenuated with the cold war whose fighting kept them alive. Gorbachev's common European home might then be a collection of smaller communities, a kind of Switzerland writ large. Maybe that is what he is driving at even in the Soviet Union. □



Uranium bites dust

The British government seems bent on letting past investments in nuclear power be wasted.

THE British seem to have a flair for indecision from which even the present decision-prone government is not immune. The failure to decide what should happen to the goods and people delivered to South-East England by the Channel Tunnel a few years from now would be a joke were it not likely to become a reality. Now, a modest programme of nuclear construction has been abandoned after a decade's costly effort by engineers, bureaucrats and lawyers, but for reasons not strictly concerned with the safety of nuclear power or with long-term economics. It is as if the British government shares with its predecessors a liking for the sensation of sand running through its fingers.

Ten years ago, the government came to office with a plan for building nuclear reactors to the Westinghouse pressurized-water design at the rate of roughly one a year. Construction of the first generating station (with two reactors) began two years ago at Sizewell in Suffolk, on the North Sea coast. The project has been enormously delayed by the decision that a public planning inquiry should be a 'generic' inquiry, dealing with general matters as well as the suitability of the site. The hope was to save time later in the construction programme. But the second public inquiry, in respect of a site at Hinckley Point on the Somerset coast (which closed last week) has been almost as long and painstaking. Indeed, the inquiry outlasted the nuclear programme itself, which was formally put on ice last month.

The explanation is simple, but not straightforward. The British nuclear programme has been a casualty of the government's determination to sell the electricity industry to private investors. The first stages of this huge exercise will begin next year, with the privatization of the electricity distribution companies. The sale of the generating plants, which will follow in 1991, has always seemed more complicated. Originally, it was planned that there should be two generating companies (one large and one small called National Power and Powergen respectively) to simulate competition. Now there is to be a third, whose plant will consist of all the nuclear stations operating or being built. It may just be on the point of going out of business when the greenhouse effect prompts another British nuclear programme.

The origins of this development are not obscure, but concern the high costs of operating a series of unsuccessful and uneconomic reactors called AGRs (for "Advanced Gas Cooled") as well as the estimated costs of decommissioning a previous generation of reactors. National Power, the larger of the generating companies which was to have had responsibility for nuclear generation, may now kick itself for have lobbied so successfully to shield its future shareholders from hardship by

frightening the government with worst-case costs.

The effects of the decision are nevertheless curious. The new nuclear company will be cast in the mould of the Atomic Energy Authority created in the 1950s as the spear-carrier of the new nuclear industry. Over 40 years, the authority has consumed a substantial fraction of Britain's engineering skill, together with substantial amounts of money. It remains a technically excellent organization, but without a nuclear programme, it might as well be renamed (its Harwell laboratory has become an excellent engineering and materials laboratory) or even be wound up when the fast reactor programme in Scotland comes to an end. Future British governments may find it simpler, on future occasions, just to dig holes in the ground and fill them in again. □

Generating physicists

A British minister hopes that compulsory physics will make more physicists, but he could be wrong.

STUNG by the example of the Japanese, all industrialized states are these days worrying about the sufficiency of the supply of technical people, scientists and engineers in particular. Generally speaking, there are too few of them. Generally speaking, the poorer industrialized economies are the most deprived. There is no accident in that: science education costs time and money, at least in comparison with the language and literature of the region (unless, ironically, the language is Japanese). Even so, it is curious that one of the British government's least celebrated innovations of the past few years is the introduction of a national curriculum for primary and secondary schools, partly with the objective of increasing the supply of scientists and engineers. How will it work out?

To judge from what Mr Alan Howarth, a minister at the British Department of Education and Science, was saying the other day, not much has changed. Speaking to the Institute of Physics, Howarth said last month that the national curriculum, by introducing students to physics at an early age, must eventually increase the supply. But the truth could easily be quite different. Past experience in British schools with the old curriculum, allowing for specialization at too early an age, shows that there is nothing like a poor teacher to drive young people into other pursuits. Howarth was promising more teachers the other day, but will they be good teachers? And if not, may it not be the worse for science education that young people should be forced to sit through classes that will convince them that they should opt for the arts and humanities? □

If Howarth really had the supply of technical people at heart, he would settle instead for a compulsory curriculum that did hardly more than persuade young people that these technical fields are interesting and important, letting them learn to be professionals when they are ready for the task. □

Back to the drawing board for Yucca Mountain repository

- DoE admits research was inadequate
- Government to sue Nevada for right to dump?

Boston

PLANS by the US Department of Energy (DoE) to build the nation's first and only high-level nuclear waste repository were set back to square one last week when a top-ranking DoE official unexpectedly acknowledged that the department lacks confidence in its research on the facility and will completely revise the project. The announcement effectively scraps the \$500-million-worth of research on the repository site, at Yucca Mountain in Nevada, that the DoE has conducted over the past two years. And the sudden change of heart throws the future of the high-level waste dump into uncertainty, a position that is exacerbated by an imminent legal showdown between the federal government and the state of Nevada.

Disclosing the overhaul of the programme at a nuclear industry convention in San Francisco, Deputy DoE Secretary W. Henson Moore said that even with "optimistic assumptions" the new changes will cause a delay of at least seven years in

the project. As outlined by Moore, the department's revised timetable means that the repository will not open until the year 2010 at the earliest — 12 years beyond the 1998 deadline imposed originally by Congress when it commissioned the high-level waste project in 1982.

Further clouding the facility's fate is the fact that the congressional instruction to the DoE is not to build the repository, but merely to assess the suitability of the proposed site, which lies deep in the volcanic rock of Yucca Mountain, 100 miles northwest of Las Vegas, Nevada. The new schedule pushes even this approval process beyond the year 2000, and if Yucca Mountain were then deemed unsuitable, the search for a suitable repository site would have to start anew.

The DoE's research on the site until now has been roundly criticized by the Nuclear Regulatory Commission and other agencies, as well as by some of the DoE's own scientists. Some observers have praised the DoE for its candour in

repudiating its previous work. The DoE has now promised that future research on the facility will be subjected to review by the National Academy of Sciences, by the state of Nevada and by industry.

But news of the programme's setback has been greeted with dismay in the nuclear industry. The cooling ponds which between them hold hundreds of thousands of spent fuel rods from the country's 110 nuclear power plants are filling rapidly; many are nearing capacity, and some plants have already been forced to build new temporary waste facilities to allow operations to continue.

US electricity utilities have already raised some \$3,000 million from rate-payers for the construction of a waste repository. More than one sixth of this money has already been spent on the now-abandoned research. The DoE, meanwhile, is bound by contracts to begin accepting waste from the nation's electric utilities in 1998, and the utilities, having fulfilled their side of the bargain, balk at the idea of spending more of their money to build temporary storage facilities. The DoE is expected to ask Congress for permission to build an interim storage facility in addition to the Nevada site.

The state of California has already passed a law prohibiting construction of any new nuclear power plants until the issue of nuclear-waste disposal has been settled. Some industry representatives have already begun to step in to this tangle of policy: the Electric Power Research Institute, for example, a research group representing the nation's electric utilities, is said to have asked for a role in overseeing the work on the repository.

Most observers believe that even the DoE's revised timetable is extremely optimistic. Complicating matters further is the resounding rejection of the site by the state of Nevada. In two bills passed last spring, state legislators exercised what they believe is their right to veto the Yucca Mountain site. But Moore said the Department of Justice will sue the state of Nevada to force it to accept the site.

Nevada legislators, including Governor Bob Miller, say they welcome the legal challenge. They have opposed the process unflinchingly ever since they were singled out as the only site under consideration by a surprise congressional move in 1987.

Most recently, the state has denied the DoE a permit to dig an exploratory shaft at the Yucca Mountain site. Robert Loux, Miller's nuclear-waste policy coordinator, says he believes there are at least three separate issues in the case that could ultimately go to the Supreme Court. As Loux sees it, the matter hangs on a simple question of state's rights. "Shouldn't a state", he asks, "have a fundamental right to preserve its economy and environment in the face of a noxious, unwanted federal project?"

Seth Shulman

HUMAN FRONTIER PROGRAMME

Switzerland knocking on the door

Tokyo

JAPAN's Human Frontiers Science Programme, which opened its European office in Strasbourg on 21 November (see *Nature* 342, 334; 1989), is already attracting the interest of countries not among the present participants. Switzerland has offered support to the tune of \$150,000 per year, and Sweden and Australia are also keen to join the programme.

Switzerland's possible membership was discussed at a meeting of the Human Frontiers board of trustees in Strasbourg the day before the office opened. Although the Human Frontier programme is underwritten by the Japanese government, the board of trustees consists of representatives from the G-7 group of Western nations (Japan, United States, Canada, West Germany, Great Britain, France and Italy) and the European Community (see *Nature* 339, 573; 1989). When Strasbourg was chosen as the European home of the Frontier programme, France made a financial commitment to the effort, and Switzerland has offered to pay for Frontier fellowships and to provide some financial support for grants and workshops. In return, the Swiss want a place on the council of scientists that sets research goals for the programme. At present, each of the G-7

nations has two scientists on the council.

Some of the seven, including Japan, were keen to accept Switzerland's proposal immediately, but representatives of other nations felt that rules for accepting new members should first be established. But Toichi Sakata, director of the Frontier office of the Science and Technology Agency, is confident that Switzerland will gain membership after the next trustees' meeting in April. Sweden and Australia have also recently made proposals for membership similar to Switzerland's, Sakata says.

The meeting included "tough" negotiations over the secretariat that will manage the programme in Strasbourg. Several nations were "aggressive" in their desire to get secretariat positions, says Sakata. But he takes this as a positive sign of interest. As well as appointing Sir James Gowans of Great Britain as secretary general (see *Nature* 342, 463; 1989), the meeting also created permanent positions on the secretariat for France, West Germany and Japan. Other nations will contribute experts as needed. The board of trustees also approved nominees for the council of scientists (European Community members have yet to be nominated).

David Swinbanks

How the earth moved . . .

San Francisco

ALTHOUGH Santa Cruz was the city closest to the epicentre of California's Loma Prieta earthquake on 17 October, the damage was worse in west Oakland and parts of San Francisco, a consequence of the way that complex local geology transmits the earthquake's energy. After travelling around the entire Bay Area to survey the degree of ground motion, two teams of researchers from the United States Geological Survey (USGS) in Golden, Colorado, have made a preliminary estimate of the magnitude of the Loma Prieta earthquake according to the Modified Mercalli intensity scale, a qualitative measure of ground shaking, rather than the familiar Richter scale, which measures the energy released at the epicentre.

The Mercalli scale runs from I to XII, with magnitude I being imperceptible and magnitude XII corresponding to complete destruction. A rating of VIII, which was assessed by the USGS researchers for Santa Cruz and its immediate vicinity, is characterized as the threshold of structural damage to buildings — wood-framed houses move on their foundations and unreinforced masonry walls tumble to the ground. But the localities that suffered the greatest destruction, west Oakland (where the Cypress section of Interstate highway 800 collapsed) and the Marina and Market Street areas of San Francisco, were assigned a IX rating.

Researchers suspect that thick deposits of alluvium near these heavily damaged areas amplified seismic waves. Alluvial layers of unconsolidated sediments and mud ring the perimeter of San Francisco Bay up to 200 feet thick, responding like "a bowl of jelly", absorbing and amplifying the seismic waves passing through the rock beneath, according to USGS geophysicist Paul Thenhaus.

As a result, he said, "it's not unusual to get higher ground motions at long distances from the epicentre in a large earthquake". Indeed, similar circumstances beset the 8.1 magnitude (Richter) earthquake on 19 September 1985 that devastated large parts of Mexico City. Even though the earthquake's epicentre was 350 km away, the city suffered widespread damage because its central and eastern portions are largely on an old lakebed.

The Mercalli rating does not provide a completely clear picture of the Loma Prieta event. Single-storey buildings in Santa Cruz fared much worse than those in San Francisco and Oakland, although the latter were shaken harder. The discrepancy results from differences in propagation between low- and high-frequency seismic waves, and from varying resonances of different kinds of building (see *Nature* 341, 676; 1989). Also clouding

the analysis was the liquefaction of sandy soils, particularly in the San Francisco Marina district, which led to serious damage even though the ground motion was relatively small in that area. To corroborate their assessments, the USGS teams studied areas of similar damage that appeared to be uninfluenced by such ground effects.

In addition to their direct observations, the researchers incorporated results from

1,313 questionnaires, distributed to postmasters and police and fire stations around California and in parts of western Nevada and southern Oregon, asking for eye-witness accounts of the effects of the earthquake. Not all the forms, which include 63 questions with up to seven different answers each, have been returned. Carl Stover, head of the US Earthquakes Project for the USGS, does not expect a final report, incorporating data from research papers and other sources, to be complete for two years.

Robert Buderl

. . . and the pollution didn't

San Francisco

WHEN the San Francisco Bay Bridge, a section of which collapsed during last month's earthquake, reopened on 17 November after its month-long closure, local air-quality officials at first reported that the shutdown, which forced many commuters to leave their cars at home and travel instead by ferries and trains, had led to a measurable fall in air pollution from automobile emissions.

But now officials of the Bay Area Air Quality Management District (BAAQMD) admit they may have spoken too soon. An array of complications, including weather and heavier-than-normal car traffic on other bridges, seems to have made doubtful the hoped-for 'case study' of what would happen if people did not use the Bay Bridge.

Researchers are now saying that they have no scientific evidence for cleaner air during the bridge closure, a great disappointment to BAAQMD's board of directors, which had hoped to drive home to the public the benefits of staying off the roads.

The San Francisco Bay Area, with some 5.5 million residents, is the fourth-largest US metropolitan area. Before the earthquake, 243,000 vehicles crossed the San Francisco Bay Bridge daily, and the area's four other main bridges carried a total of 273,000 vehicles. These cars, trucks and buses contributed half of the 500 tons of hydrocarbons released daily into the atmosphere, and 90 per cent of the 3,000 tons of carbon monoxide. Of the two, hydrocarbons, which are precursors to the formation of ozone and thence of smog, are regarded by air-quality officials as the more serious pollutant.

At first, BAAQMD officials simply removed the Bay Bridge vehicles from their calculations, and guessed there should have been a 6.7 per cent reduction in hydrocarbon and carbon monoxide levels while the bridge was closed. But figures collected at four alternative bridges showed their traffic went up to 435,000 vehicles a day during the Bay

Bridge closure, making up two-thirds of the difference. Moreover, the extra traffic caused heavy congestion, and according to James Sandberg, section chief of meteorology and data analysis for the air-quality district, in slow-moving traffic the emission per car goes up.

Added to this complication is the fact that ozone forms only when the weather is warm enough to provide the radiant energy to drive the needed photochemical reactions. And during hot weather an inversion layer of warm air sits over the surface, trapping and concentrating pollutants.

In the moderate Bay Area climate, ozone levels have never exceeded federal standards from mid-October until early spring. "It's just not our ozone season", said Milton Feldstein, BAAQMD's air-pollution control officer.

Even if all the cars in the area had stayed off the roads, there probably would not have been a measurable change in ozone levels.

"I hate to say maybe we should have had the earthquake in September when it was nice and warm, because they'll lynch me", said Feldstein.

The end result is that ozone levels during the Bay Bridge closure were down slightly, but did not go beyond the bounds of normal seasonal variations, making any useful scientific conclusion impossible. But that does not mean the unique case study of the failed bridge is a total loss. Researchers are now trying to determine more exactly the number of cars on the road in the month following the earthquake.

Taking into account the slower speeds caused by increased congestion, they will then try to make a more accurate projection of emissions reductions and create a computer model to see if the decrease in emissions would have made a difference in ozone season. Feldstein said that BAAQMD might then still obtain evidence to convince people of the virtue of staying off the road.

Robert Buderl

Public debates on ethics

Washington

BEGINNING in 1991, town meetings may be held throughout the United States to inform the general public about the human genome initiative and to solicit opinions on the ethical, social and legal issues it raises. This is one of the recommendations of the human genome advisory committee's ethics working group, which presented its first report this week at a meeting held by the full committee to discuss the latest five-year plan for the human genome initiative. The new plan does little more than reiterate the goals outlined previously, but the committee stresses with more urgency that completing the project within 15 years, as planned, will require a doubling of the annual budget to \$200 million within the next two years.

The possible social perils inherent in the human genome programme, such as loss of employment or insurance, are thought to be most likely to occur in the interim phase when genetic diseases can be detected but not treated. By establishing a programme to tackle ethical, legal and social issues long before most of the scientific results start to come in, the two agencies responsible for the initiative, the National Institutes of Health (NIH) and the Department of Energy (DOE), hope to anticipate and avoid problems. Of the budget for human genome research, 3 per cent is being allocated to the study of these social issues.

The ethics working group, chaired by Nancy Wexler of Columbia University, recommends that research into issues such as the ownership and control of genetic information, genetic counselling and the commercialization of results from the human genome initiative should be examined in a series of 'focus groups' which would include representatives of insurance companies, industry, unions, the media and lawyers as well as geneticists and ethicists. With the information thus amassed, the working group would formulate policies to put before the human genome advisory committee. Plans for the first workshop are under way.

The working group also calls for collaboration between NIH and the National Science Foundation to develop educational material on the human genome initiative, aimed at groups such as students, the media, medical practitioners, researchers and recipients of genetic services. It also says there should be postdoctoral fellowships for biomedical researchers who want to study the ethical, legal and social aspects of human genome research; technical education should also be made available to lawyers and ethicists. In all these activities, the working group says the United States

should collaborate with other countries, including those that are not directly involved in the human genome initiative. HUGO (the international human genome organization) or UNESCO could co-ordinate international efforts, the report suggests.

Also before the full advisory committee this week is a proposal to set up a body with responsibility for establishing a national informatics programme. This body, to be called the joint informatics task force, would consist of experts in molecular biology or computer science from academic, government and industrial laboratories and would be responsible for addressing the technical problems of handling the mass of data that the human genome initiative will generate.

The proposal for a joint body comes from the informatics groups of NIH and the DOE, which would cede their roles to the new body. These groups define the immediate goal of the task force as the creation of a database that provides easy access to current physical mapping, genetic mapping, chromosome mapping and sequencing information. The task force would also be responsible for developing software to support large-scale mapping and sequencing projects. Differ-

ent groups will develop their own databases and software, but the aim of a coordinating committee is to avoid "a costly reinvention of the same wheels over and over again at different labs".

The task force will also encourage communication between computer scientists and biologists working on the human genome initiative. The problems that arise from it "will not be solved merely by biologists describing their problems in biological terms and hoping that computer scientists will be able to provide a solution", says the report. Special funding programmes should be established, it says, to provide graduate and postgraduate training in biology to computer scientists and vice versa.

Workshops, short training courses and long-term exchanges of personnel could also improve collaboration.

The more controversial items on the agenda of the new task force would include deciding who should have access to data and how to charge private industry and foundations for access to academic and government databases. It will also examine who will have responsibility for maintaining genomic databases, what level of funding the informatics component of the genome project should have and how to collaborate with the international community.

Christine McGourty

BRAZIL

Changes for science, but for how long?

São Paulo

A YEAR of perplexing reorganization for Brazil's science administration was topped off last week with the re-establishment of the Ministry of Science and Technology by President José Sarney. The ministry had been abolished by Sarney just 11 months earlier. In the meantime, public and congressional protest had led to the reappearance of the ministry as the Special Secretariat for Science and Technology (see *Nature* 342, 371; 1989). The latest move puts the ministry exactly back where it started at the beginning of the year. The special secretary, Décio Leal de Zagottis, remains at his post but changes his title to "Minister".

"It is a joke in bad taste", was the reaction of Ennio Candotti, president of the Brazilian Society for the Progress of Science (SBPC), who questioned the timing of the change. Brazil will vote to choose a new president this month, meaning that Zagottis will run the new ministry only until 15 March when the new president officially takes office.

The decree that re-establishes the ministry explains that "science and technology matters are vital for the socioeconomic development of the country", and that experience now shows the minis-

try is needed. According to the decree, the existence of the ministry will make it easier to deal with international institutions and other agencies in Brazil.

But scientists are baffled by the sudden reorganization. Candotti says it is pointless just to change the name to 'Ministry' when it is lack of resources that is the real problem for science in Brazil. The proposed budget will provide only \$100 million for science in the first half of 1990 although \$400 million is needed "barely to survive", he says. FINEP, the Agency for Financing Studies and Projects which provides support for both basic research and technological development in industry, is particularly hard hit in the new budget. "The budget for 1990 is ridiculously small", says Luiz Pinguelli Rosa, head of the Federal University of Rio de Janeiro's distinguished graduate school of engineering. He says that scientists are not against the ministry but are astonished by its re-appearance so soon before the election.

The constant change and confusion leads many scientists to view Sarney's government as just a succession of administrative disasters. It shows "Sarney's lack of administrative seriousness", said Rogério César de Cerqueira Leite, of the University of Campinas. Ricardo Bonalume Neto

Copyright vs patents for software

Washington

DISAGREEMENTS within the computer science community over the best way to protect software surfaced last week at a discussion meeting organized by the computer science and technology board of the National Academy of Sciences. Software has traditionally been protected through copyright laws, but researchers are increasingly taking out patents on their programmes, and are divided over the issue of whether patent protection should be strengthened or weakened.

Among the problems created by the recourse to patent law is the delay in processing applications at the US Patent and Trademark Office (PTO), and the ambiguity of some patents issued, which has on occasion erupted into expensive law suits. The PTO is also causing irritation among academics and industrialists in the computer science community by issuing patents for software that many consider too unoriginal to merit protection. Robert Spinrad of the Xerox Corporation complained that the confusion was having a "stultifying and dulling effect" on the industry. Dan Brinklin, now president of Software Garden Inc., who was forced to sell his former company after a successful but expensive patent dispute, said that the software industry is "very scared" by the patent process.

The debate centred upon what would best serve the public interest. John Schoch of Asset Management Company, a venture capital group based in California, argued strongly that patent protection must be strengthened as this encourages innovation and therefore benefits the public. But Anita Jones of the department of computer science of the University of Virginia argued that patent protection should be weakened because the public benefit lies chiefly in the unrestricted dissemination of information.

There was also disagreement over the use of copyright to protect software. Although some complain about the copyright system, it has the advantage of being enforceable worldwide through international copyright agreements.

Howard Figueroa, vice-president of commercial and industry relations at IBM Corporation, was among those who spoke strongly in favour of the current system which relies heavily on the use of copyright. His response to heated discussion between those who favour strengthening patent protection and those who would like to see it weakened was "if it ain't broke don't break it".

The 20–25 per cent annual growth in the computer software industry in the past 10 years proves that the current system is working, and the problems could easily be put right he argued.

Christine McGourty

Late award reveals problems

Tokyo

AFTER waiting for three decades, the US electronics giant Texas Instruments (TI) has been awarded a Japanese patent for the technology that forms the basis of nearly all integrated circuits. The patent is expected to bring in hundreds of millions of dollars in royalties and will provide a powerful bargaining tool for TI in renegotiations of cross-licensing agreements with Japan's electronics companies. But although TI has hailed the agreement as a new sign of respect for intellectual property rights in Japan, the prolonged process of patent application only highlights the problems that foreign companies face in protecting their rights in Japan.

The newly patented invention goes back to 1958, when Jack Kilby, a TI engineer, showed how to build transistors, capacitors and resistors together on a semiconducting substrate to make an integrated circuit. A US patent was applied for in 1959 and awarded a few years later. A Japanese patent application was filed on 6 February 1960. But the Japanese patent office objected to the broadness of the application, which was then split into several smaller claims.

But still objections were raised and the proceedings delayed, and only on 27 November 1986 was the central part of the original patent application accepted and published by the patent office.

Many of Japan's electronics companies immediately filed opposition proceedings but finally, on 30 October this year, the patent was granted.

The Nihon Keizai Shimbun, a financial newspaper, reported the patent award two weeks ago and predicted that it could bring TI royalties of as much as ¥80,000 million (\$550 million) a year.

But most of Japan's semiconductor manufacturers already have cross-licensing agreements with TI, and they are expected to trade other technology rights, rather than pay royalties, when their agreements with TI come up for renewal in the next year or two.

But smaller manufacturers, as well as newcomers to the semiconductor trade, which do not have leading-edge technology to barter with, are expected to have to pay out large amounts in royalties, according to analysts.

Masahiro Miki, a spokesman for TI Japan, says the patent award shows the importance that Japan is now attaching to intellectual property rights. And he points to TI's success in suing Japanese and Korean companies in 1987 for violations of the company's patent rights for dynamic random access memory (DRAM) technology as another good men. But he refused to discuss why the Japanese patent office took more than 29 years to process

the Kilby patent.

Compared to US patents, Japanese patents tend to be much more narrowly defined, each covering minute aspects of an innovation.

The Japanese patent office is flooded with about half a million patent applications a year, which take on average 5–7 years to process. In the United States, by contrast, 200,000 applications come in every year, and most are dealt with in 18 months.

According to Mitsuru Miyata, editor of the newsletter *Nikkei Biotechnology*, the Japanese patent system is designed not for the protection of intellectual property rights but rather for the import and use of foreign technology. Japanese companies seldom fight each other over patent rights, he says: instead, competitors go out of their way to file narrow patent applications that differ from each other in only some minor respect, allowing them to compete commercially without violating each other's patents.

When a foreign company tries to protect a broad technological invention in Japan, it is forced to break down the patent into a series of narrow applications, each of which has to fight its way through the system. The delays provide a crucial advantage for Japanese companies. If TI had been awarded its patent 5 or 10 years ago, it would have been a far greater financial burden for Japan's electronics giants, because they then lacked the sophisticated technology then they use in cross-licensing agreements with the US company.

This is not the first time the Kilby patent has been at the centre of confrontation between the United States and Japan over technology. In the 1960s, the Ministry of International Trade and Industry (MITI) rejected initial efforts by TI to set up operations in Japan, according to David Lammers of the Tokyo office of the *Electronic Engineering Times*. But the US company suggested that any Japanese integrated circuit incorporating the Kilby patent would be barred entry to the US market and so MITI backed down, Lammers says.

Although TI says the award of its patent is a sign of change for the better in Japan, the irony is that under present Japanese law the patent would never have been granted. Under current rules, a patent application expires 20 years after the application is filed or 15 years after publication, whichever comes first, says Yasunori Ohtsuka, a patent lawyer in Tokyo. But the TI application in 1960 was submitted just two months before the current law came into effect, and under the old law, which still applies to the US company's case, there is no time limit.

David Swinbanks

Science budget falling behind

Cape Town

RESEARCH prospects in South Africa are looking bleak. Although next year's budget has yet to receive formal approval, the Council for Scientific and Industrial Research (CSIR) is expecting an increase of at most 4 per cent. With inflation running at 15 per cent, and research costs hit further by the weak rand, significant cut-backs are inevitable. Even so, science and education seem to be relatively well-treated by the de Klerk government, which is making a concerted effort to curb state spending; a 15 per cent cut in the defence budget, for example, has been suggested.

Worst affected has been the CSIR's Foundation for Research Development (FRD), whose responsibility its head, Rhein Arndt, describes as "manpower development in science and technology".

This includes the funding of university research and the provision of student grants. Last year, the CSIR's national

SPACE AGENCY

Arianespace forced to twiddle thumbs

Paris

TECHNICAL problems, small print and the San Francisco earthquake have left Ariane-space, Europe's commercial satellite launch consortium, with nothing to launch this month, despite a full order book and a tight schedule.

Originally, the French SPOT 2 commercial Earth observation satellite — the overdue replacement for the successful but ageing SPOT 1 — was scheduled for launch last month. But 'technical problems' with the Ariane 40 rocket led Ariane-space to delay the Flight 35 launch until December, pending further tests.

However, the Superbird A satellite, built by Ford Aerospace for the Japanese, was also scheduled for a December launch and a postponement would have incurred penalties for Ariane-space.

So the consortium asked the French national space research centre (CNES), which owns SPOT, to let the Japanese satellite go up first.

Because Spot 1 is still sending pictures, Spot 2 was rescheduled for launch on 10 January. In the meantime, Superbird A, which was undergoing final tests at Ford Aerospace in Palo Alto, was shaken by the San Francisco earthquake, and the manufacturers decided it would be unwise to send the satellite to the Kourou launch site without more tests. As Spot 2 could not be taken out of wraps again at short notice, this has left Ariane-space with two launchers ready to go, but no payload.

Peter Coles

laboratories were taken from FRD and given to another arm of CSIR, Research, Development and Implementation (RDI) (see *Nature* 332, 197; 1988). The aim was to move towards 'near-market' research, and the RDI has been successful in supporting research with contracts from government departments and the private sector. Last year, 38.8 per cent of the CSIR's income was derived from outside contracts, and this is expected to rise to 41.5 per cent this year. In compensation for this move, the FRD expected to get more of the CSIR's overall budget as the RDI became more self-financing. But this has not happened: this year, FRD was allocated the same proportion of total government support as last, and no change is expected in the coming year. The net effect has therefore been a shift of money from basic to applied research.

Although the CSIR's grant is negotiated from the Treasury by the Ministry of Trade, Industry and Tourism, the RDI and the FRD budgets are considered to be separate entities, and are dictated by the government. The cabinet is currently considering a request by the FRD to be moved to the national education budget, which is in the hands of Minister Gerrit Viljoen. He is more enthusiastic about basic research, and also seems to have the ear of President F. W. de Klerk.

The FRD also runs the country's four national research facilities: the South African Astronomical Observatory, the Radio Astronomy Observatory, the Magnetic Observatory and the National Accelerator Centre (NAC). Last year, NAC alone cost CSIR R30 million (\$12 million), 34 per cent of the CSIR budget, and many of the country's scientists feel that NAC is an under-used and unaffordable luxury.

Last month, FRD initiated an independent review, by scientists from outside South Africa of the quality of work being carried out at NAC, its impact on the development of technical expertise in South Africa, and NAC's place in the country's science system. Arndt says that the future of NAC is "something that the science community will have to take a hard decision on".

At the root of all these problems is the small proportion of gross domestic product (GDP) that South Africa spends on research and development. In 1987, research and development spending from all sources amounted to 1 per cent of GDP (R1,700 million, or \$700 million), with roughly equal amounts coming from government and the private sector. This is a much smaller proportion than is spent by Britain or France (2.3 per cent of GDP in 1987), and is less even than India (1.1 per cent).

Michael Cherry

Technical studies to be compulsory

Sydney

AN ambitious reform of the education system in New South Wales will make the study of mathematics, science and technology compulsory for high school students in the province. Under the present system, the only compulsory subject for final year students taking their Higher School Certificate is English, but by 1992, following the introduction of the government's white paper (policy document) on curriculum reform, students will also have to take at least one technical course.

Until now teachers have been providing courses in technology and computing, of a mostly vocational nature, under a category known as Other Approved Studies (OAS). The government intends to reduce the OAS from 10,000 to 500 subjects, substituting new technology courses instead. The white paper will also make it compulsory for 13–15 year-olds to study a language and a technology course. In other changes to the curriculum, courses involving AIDS and drug education will be integrated into a new mandatory health and physical education programme.

The government acknowledges that the major difficulty in the introduction of these reforms is a shortage of teachers of technology, science and mathematics. A spokesperson for the Minister for Education said that the department would be negotiating with the NSW Teachers' Federation for a salary increase for staff in these fields. But the teachers are unlikely to go along with such changes. Sue Simpson, spokesperson for the NSW Teachers' Federation, said she found the suggestion of differential rates of pay surprising, and argued instead for "an overall increase in teachers' salaries to entice teachers back into the system".

Tania Ewing

ASTRONOMY FROM SPACE

Solar max falls at last

Washington

THE Solar Maximum Mission satellite (SMM), launched in 1980 to observe flares and other activity during the peak of the last solar cycle, almost survived until the peak of the present solar cycle, expected in about four months. But last Sunday morning, 3 December, it re-entered the Earth's atmosphere and burned up over the Indian Ocean near Sri Lanka. In 1984, SMM became the first satellite to be successfully retrieved, repaired and returned to orbit by the crew of a space shuttle mission. SMM's swansong was its observation of gamma-rays from Supernova 1987A, which provided the first observational confirmation that supernovae remnants are powered for the first year or so by the radioactive decay of ^{60}Ni synthesized in the initial explosion (see *Nature* 331, 416; 1989). David Lindley

Grappling with dirty past

Munich

THE East German environmental movement took its first shaky steps toward political responsibility last week with the formation of the 'Green Party of the German Democratic Republic'. In its first statement, the party criticized materialism and advocated reducing environmental pollution.

But the party, founded over the objections of other ecologically minded people in East Germany, will have to develop a more solid programme in the next few months if it is to achieve even a fraction of its goal of ecological reform in a socialist framework. Nevertheless, the founding of the party was considered significant enough to be reported widely in the East German media and to attract support from the West German Green Party.

Co-founder Gerhard Bächer admits that winning support will be difficult in a country where everyone pays lip service to ecology. Even members of the ruling Communist Party have called for an "ecological reconstruction" of the country. Bächer says that the Greens have received anonymous phone calls from Communist Party members, insinuating that they have shown right-wing tendencies.

A bigger threat is the fragmentation of green thought between the new party and other left-wing groups such as New Forum. Bächer claims that the difficulties have been dealt with, but a West German Green who visited last week said that many disagreements remain. Members of the opposition movement New Forum also fear that they could be split by difficult choices between preserving the environment in East Germany or opting for profitable but polluting activities that bring in hard currency. The environmental side had been gaining ground in East Berlin earlier this year, but the collapse of the government and the need to earn Western money may turn the tide.

If it can arrive at a common platform, the Green Party will have a lot to do in East Germany, which has always relied for energy on soft bituminous coal. Huge strip-mining machines have churned up the countryside, and when the coal is burned, sulphur compounds are released directly into the atmosphere. Smog is a permanent problem in all East German cities, made worse by the black, foul-smelling exhaust from the two-stroke engines of Trabants, the least expensive East German cars.

The East German Greens are probably the only Green Party in the world that favours nuclear energy, but they support it only for want of alternatives. But their support does not extend beyond completing the nuclear power plant at Stendahl that is under construction. They are against

new plants, but can suggest no likely alternatives that East Germany can afford.

Although East Germany has never had a significant nuclear accident, the uranium mines in the south of the country have been a continual health threat for East Germans who live nearby. The East German government has begun to publish radiation and air pollution figures for the first time in its history.

The East German Greens will be preoccupied with such basic issues for a long time before they need to consider genetic engineering and other technologies that

IN VITRO FERTILISATION

Reformation of advisory board urged

Washington

THE US Congress can now be added to the list of agencies trying to prod the US federal government into reconstituting the Ethics Advisory Board (EAB), the committee with responsibility for overseeing research on *in vitro* fertilization (IVF). In a report issued this week, the House of Representatives committee on government operations urges the government to abide by the regulations of the Department of Health and Human Services which require it to set up the board. It is an "embarrassment" to the department and to the Public Health Service that the last five secretaries for health have ignored these regulations, says the committee.

It also urges the present Secretary for Health, Louis Sullivan, to implement the board's recommendation, made in 1979 before its charter was allowed to expire, to exempt research on embryos less than 14 days old from the need to be reviewed by the EAB. Other restrictions would be adequate, such as requiring the informed consent of the couple from whom the embryo was generated or restricting the number of eggs that can be fertilized.

Reinforcing the comments made by the Institute of Medicine in a report issued in November (see *Nature* 342, 218; 1989), the committee complains of the inadequacy of current IVF research and the poor infertility services available.

In line with the controversy surrounding this issue, five of the 15 Republican members of the committee signed a dissenting statement which says that there is no evidence that significant increases in federal funding for this research will improve its quality. It also says that there are fears that an increase in funds for infertility research could strip funds from AIDS research, and that, because the private sector is involved in this research, this is "not the best use of scarce tax dollars".

Christine McGourty

exercise the West German Greens.

What may be the biggest challenge for the Greens in East Germany has little to do with ecology. Bächer warned against a "disturbing" tendency to favour reunification with West Germany as a way to help East Germany out of economic decay, saying that the West German style of economic success has its consequences, not least for the environment. "People's eyes popped out" on their trips to the West during the first weeks after the opening of the wall, said Bächer, but he and his party hope to convince them that they should consider the price before buying the goods.

Steven Dickman

COLD FUSION

An old-fashioned love-song

Tokyo

Cold fusion reared its head again last week in Japan with two independent research groups claiming to have detected strong neutron bursts from electrochemical experiments. The news was splashed across the front pages of newspapers, reported on national television and picked up by the US *Wall Street Journal*. But Japanese experts on neutron monitoring are sceptical.

Kunihide Nishizawa and Nobuhiko Wada of Nagoya University say they detected neutron emissions 20,000 times background levels from palladium electrodes that had been soaked in deuterium gas then subjected to a high voltage. Close on the heels of this announcement, a group at Osaka University, headed by Yoshiaki Arata, claimed to have detected neutron emissions 25,000 to 2.5 million times background levels when a voltage was applied to thick palladium rods immersed in heavy water.

But, according to one Japanese expert, the detector used by both groups, called a BF3 counter, is intrinsically very sensitive to vibration, and sometimes shows pulse-like signals unrelated to neutron emission. Wada and Nishizawa should present pulse-height data before they claim neutron detection, he says.

The Osaka University results are more highly regarded, because multiple detectors were used and pulse-height data were presented. But neither group has examined the energy spectra of the neutrons and thus cannot claim that fusion is the cause, experts say. As one researcher from Japan's new National Institute for Fusion Science put it "we want evidence that the signals arise from neutrons due to nuclear fusion processes" but "some Japanese scientists still sing an old-fashioned love song composed by Pons and Fleischmann".

David Swinbanks

SPACE EXPLORATION

NASA outlines path to Mars

Washington

PRESIDENT George Bush's vision of lunar and martian exploration, made public on the twentieth anniversary of Apollo 11's Moon landing (see *Nature* 340, 249; 1989), is so far no more than a vision. Trying to inject some reality into these grand schemes, the National Aeronautics and Space Administration (NASA) last month released a report that offers five exploration strategies.

But the report, the result of a three-month collaboration by all of NASA's divisions, confines itself to making suggestions, which come without price tags, that the president's National Space Council might wish to adopt with a view to working them up into detailed plans.

In the first strategy, the astronauts return to the Moon in 2001 and occupy a lunar base permanently in the following year; the first mission to Mars is in 2015. But this tight schedule depends on accelerating the development of the space station so that its assembly is complete by 1997, two years earlier than planned.

Financial problems have already forced NASA to plan a new 'rephased' space station programme which stretches out several parts of the programme by six months. The Shuttle-C, a new cargo-carrying launch vehicle needed to ferry space station components from Earth, is the only new vehicle that could be ready in time to meet this schedule.

In the second strategy, the first lunar mission is also planned for 2001, but the first human Mars landing is brought forward to 2011, delaying scientific research on the Moon but accelerating it on Mars. This would require the concentration of funds in the first decade of the next century so that hardware for both the lunar and the martian outposts could be developed in parallel.

NASA also outlines a third possible plan which is similar to this one, and a fourth which delays the major milestones by two to three years.

In the final strategy, all the major steps in the programme are delayed: the return to the Moon is in 2004; the first landing on

CONSERVATION

Too many cooks preserve diversity

Washington

THE preservation of biological diversity in the United States and worldwide is a project with many supporters. Paradoxically, that seems to be causing trouble. Earlier this year, the National Science Foundation (NSF) argued that, because of its long-standing support of ecological research, it should be given a role in national efforts to document and preserve plant and animal species (see *Nature* 340, 585; 1989). Meanwhile, plans have been stirring in Congress to set up a National Center for Biological Diversity with the Smithsonian Institution at its helm, but at a congressional hearing last month Robert Jenkins, vice-president for science programmes at the Nature Conservancy argued that his organization ought to play the lead role because of its past and continuing efforts in species preservation.

The purpose of the hearing, before a subcommittee of the House of Representatives committee on Merchant Marine and Fisheries, was to discuss two principal questions embodied in proposed legislation: the establishment of a national centre to co-ordinate data collection and research concerning biodiversity, and to incorporate into the charter of various federal agencies language to make species preservation a particular concern. At present, only the National Forest Service (NFS) has such language in its charter. David Wilcove, an ecologist with the Wilderness Society, argued that the NFS had not traditionally

put conservation issues high on its list of priorities, and cited the case of the northern spotted owl, which inhabits forests of the Pacific northwest: NFS forest management, Wilcove claimed, had given over such quantities of the owl's habitat to logging that the Fish and Wildlife Service, another federal agency, was now arguing for the owl to be listed as an endangered species.

It was agreed by all that the Endangered Species Act was ineffectual in preserving general biodiversity, and that some mechanism was needed to enforce conservation regulations before species were threatened to the point of extinction. But while there was general satisfaction at the hearing with the idea of a national centre run by the Smithsonian, Jenkins protested that the Nature Conservancy, a private non-profit organization, had been running such a programme on a small scale for many years.

Its Heritage Program, he said, was ready to receive federal support that would expand it into the National Center for Biological Diversity of proposed legislation, and suggested an amendment of the law that would give the Nature Conservancy the appropriate leading role. But Jenkins seemed to be in a minority of one: George Davis of the Academy of Natural Sciences of Philadelphia said there was "plenty the Nature Conservancy is not doing" and that only a suitable federal body would have the authority to enforce interagency collaboration.

David Lindley

MOLECULAR BIOLOGY

Belgium joins EMBL

Heidelberg

BELGIUM has become the fifteenth member of the European Molecular Biology Laboratory (EMBL), director Lennart Philipson announced last week. The additional support of roughly DM1.5 million (\$840,000) a year, plus an entry fee, is a modest addition to EMBL's budget, but the new funds, says Philipson, will allow EMBL to increase the number of fellowships it awards. The exact use of the money must be determined by the EMBL council. Like all member states, Belgium will contribute to EMBL in proportion to its gross domestic product. Its yearly payment will provide about 3.1 per cent of the DM49 million (\$28 million) EMBL budget.

Steven Dickman

Mars is in 2016 and a permanent Mars outpost would not be ready until 2027. Research on the Moon would begin around 2007 with Earth-controlled geological and geophysical exploration, establishment of the astronomical telescope arrays and human biomedical research in the lunar laboratory. In 2022, there would be a piloted flight to the far side of the Moon for geological exploration and to set up a radioastronomy and geophysical observatory.

To carry out this programme, NASA says it will need a substantial increase in staff numbers, from the 24,000 permanent staff it has at present. But it complains in the report about the "antiquated" federal employment structure which keeps salaries and benefits low, making it difficult to recruit high-quality scientific and technical personnel in the increasingly competitive marketplace.

NASA also wants to change the rigid budget process which allocates funds on a two-year basis, resulting in short-term planning and instability. When NASA was established it was granted "no year" funds which it did not have to spend within an established time period, and it had greater flexibility to alter programmes, to accept gifts, to buy property and to enter into contract on any terms and conditions it deemed necessary. With new legislation and regulations, these powers have been gradually eroded, but NASA says that the complexity of a 30-year lunar and Mars exploration programme necessitates change. NASA has not yet attempted a programme of this size, requiring such long-term development, and unless existing legal regulatory structures are changed, there will be "a severe penalty in schedules, resources and efficiencies" it says.

Christine McGourty

Correction

The outgoing president of the South African Council for Scientific and Industrial Research is Dr Chris Garbers (see *Nature* 342, 334; 1989).

Fair shares for Japanese

SIR—Japan's government institutions are now advertising their scheme for hundreds of foreign postdoctoral fellows (see S. S. Siddiqui, *Nature* **340**, 337; 1989 and recent advertisements in *Nature Classified*). It is good to know that our research institutions, as well as universities, have finally decided to open their doors to researchers from outside Japan. I welcome our government's decision to make its laboratories international; its efforts (including high payment) will be appreciated by foreign researchers and probably by politicians, in spite of the difficulties experienced by foreigners living in Japan.

At the same time, however, it is regrettable that non-tenure-track Japanese researchers cannot similarly join laboratories as fellows. The time is now ripe for our laboratories and universities (both now training too few post-PhD fellows) to show Japanese PhDs how they can carry out postdoctoral research in their own country. If it does not do so, Japan will miss an opportunity to keep promising young scientists wanting experience in laboratories in different fields.

YASUSHI SAOTOME

M.D. Research Co. Ltd,
760 Morooka, Kohuku,
Yokohama 222,
Japan

SIR—I read with interest the article by S. Berliner III (*Nature* **341**, 379; 1989) pointing out the difficulties of using the 2,000 characters of the Japanese language. As he recommended, the Japanese should, as a short-term measure, use bilingual signs in Japanese and English in public places, and should transcribe Japanese characters into roman letters for foreigners. As a long-term solution, however, do we really have to reform our language, which has the merits of both ideograms and phonetic signs? During my five years in New York, I found that English-speaking people overlook the fact that the language-brain interface is more important than the language-machine interface.

If a sentence written with simple characters like the alphabet is easy to understand, why is machine language composed of only 0 and 1 so difficult for humans? The brain is well suited to recognizing patterns such as faces (*Brain Res.* **342**, 91; 1985) and ideograms, and a concept is easily associated with the corresponding ideograms even in a patient with aphasia (*Cortex* **8**, 265; 1972). Japanese sentences are easy to read because they are composed of lexical portions, written with *katakana*s (phonetic signs to express foreign words) and 1,945 basic Chinese characters, and grammatical formatives with *hiragana*s (the other phonetic signs).

In fact, the circulation of newspapers in

Japan is the highest (58 per cent) in the world (20 per cent in the United States), and many attempts to persuade Japanese to read newspapers and textbooks written in English or roman-lettered Japanese have failed. The percentage of total time spent on language during elementary school is only 25 per cent in Japan but over 40 per cent in the United States, and there is almost no illiteracy in Japan. This easy Japanese education has resulted in a flourishing society.

The disadvantage of Japanese in writing has now been overcome by the widespread use of Japanese word processors and facsimile machines. Memorizing ideograms is easy for children, but not for adults. Thus, millions of biological species and chemical compounds must be written in *katakana* instead of 80,000 complicated Chinese characters (*Biochem. Educ.* **16**, 37; 1988).

People should choose a language suitable for both their brain and their culture, not for the typewriter.

YASUO KAGAWA

Department of Biochemistry,
Jichi Medical School,
Tochigi, 329-04, Japan

Don't blame us

SIR—Your leading article "Research by numbers" (*Nature* **341**, 674; 1989) quotes from a document setting out new resource management arrangements agreed last April between the Department of Education and Science (DES), the Treasury and the research councils, and wrongly suggests that these arrangements provided the basis for the government's decision not to finance the proposed national survey of sexual attitudes and behaviour.

Your readers will want to know that under these arrangements it is for the research councils to decide whether to refer matters to the DES. The initiative in seeking ministerial guidance is placed with the research councils, not with ministers.

In the case of this survey, the government decided it would not be appropriate to provide financial support from the Department of Health's research programme or, more generally, to sponsor the survey. The Economic and Social Research Council (ESRC) decided not to proceed on this basis. Contrary to your account, the ESRC was not required to refer the matter to the DES: nor did it do so.

The resource management agreement was concluded after extensive discussion between the department and all the research councils. It is designed to give greater freedom to councils in exchange

for departmental scrutiny of councils' management systems. It is not the purpose of the new arrangements to limit councils' freedom nor to intervene directly in their scientific judgements. Any such agreement would have been unacceptable to the councils which continue to operate under the well established 'arms-length' relationship with the government, embodied in the 1965 act establishing the research councils.

ROBERT JACKSON
(Parliamentary Under-
Secretary of State)

Department of Education and Science,
Elizabeth House,
York Road, London SE1 7PH, UK

Creative force?

SIR—If A. Travis (*Nature* **341**, 10; 1989) is personally having "flirtations with the untestable" and holds the belief that there is a "creative force manifesting itself in the observable universe", how does telling us this assist in furthering our understanding of nature? Surely you have received correspondence from others on more important and specific topics that better deserve column space?

The essential difference between religious beliefs such as Travis's and the "organized common sense" of science is that religion demands explanations, while science is, or should be, prepared to answer: "Nobody knows", and continue with its investigations.

Formulating unanswerable questions and then demanding answers to them is a waste of time. Formulating a theory to explain a feature of nature and then finding ways to test the theory is not. If a way is not found to test the theory, then it must be shelved until one is. A theory that life evolved here on Earth out of molecular interactions is in principle testable: for example, we can perform individual steps in the process experimentally, we can simulate it in our computers and we may one day be able actually to do the whole thing ourselves. The theory may be proved wrong, so therefore it is testable. No one has come up with a testable alternative yet, so if this one is proved wrong, we must return to "Nobody knows".

Travis looks for and finds some "unknown creative force" in this theory. That can only stem from his own religious desires for there to be one, for there is no mention of one in any of the statements of the theory that I have read.

To me, the conclusion "Nobody knows" is far more awe-inspiring and stimulating than "we can't explain this, so therefore there is a mysterious creative force in the universe", which in my opinion simply shows a lack of imagination.

K. P. APPELBY

35 Catherine Street,
Cambridge CB1 3AW, UK

The great gene shears story

John Maddox

Nature has been in hot water in Australia for supposedly misreporting the prospects for a new technique of genetic manipulation. The truth is more interesting and complicated than *Nature* (and many Australians) know.

Sydney

FOR the past month, *Nature* has been heartily reviled by many in Australia for the publication, on 12 October (341, 473; 1989), of a news report that an Australian patent application of which much was (and is still) expected had been anticipated by a patent application lodged earlier in the United States. The report has been especially hurtful to the Commonwealth Scientific and Industrial Research Organisation (CSIRO), which for the past decade, but especially in the past two years, has been under intense pressure from the federal government of Australia to win a greater commercial benefit from its substantial annual budget, but other officials of the federal government, ministers included, have also waxed indignant.

Further investigation, on a visit to Australia for other purposes, softens the assertion, in the report of 12 October, that an invention of which CSIRO, even Australia as a whole, is rightly proud has been fully anticipated by developments in the United States. To the extent that the report suggested that the Australian patent application is unambiguously pre-empted by the US patent claims, it is oversimple. But the relative status of the US and Australian claims is, to say the least, complicated and unclear. If the report had merely said (which it did) that the Australian claims will be fiercely disputed, it would have been correct.

Even so, *Nature* owes an explanation (unsolicited, as it happens) to those directly concerned — to CSIRO, to Dr Jim Peacock of the Plant Industry division of CSIRO, to Drs Wayne Gerlach and Jim Haseloff and to Mr Stephen Castle, a patent examiner at the Australian Patent Office (who was a casualty of the process of gathering the information on which the 12 October news report was based).

What follows is a circumstantial account of the origins of the disputed patents and of *Nature's* two news reports on the subject (both that of 12 October and, earlier, on 3 August, 340, 332; 1989). It is hoped that this account will be instructive (as well as interesting) for readers concerned about several features of the interaction between the scientific enterprise and the increasingly commercial environment in which it must make its way: the effects of commercialization on the open discussion of research; the extent to which publication policies are now affected by

the consideration that patent rights should not be compromised; and the intricacies of patent law, with which researchers in many fields will have to become familiar.

The relationship between research and the news media is another of the issues that arises. In some ways, *Nature's* report lacked balance. But too little help was given by some of the respondents to its enquiries (although that is hotly denied).

Ribozymes

It is a shame that the research concerned, from which both the US and Australian patent applications stem, should be one of the neatest and potentially most important applications of genetic manipulation yet. Its origins are the recognition, within this decade, that intervening sequences ("introns") may in real life (*in vivo*) be excised from freshly-transcribed RNA molecules to make functional molecules of messenger RNA (mRNA) without the intervention of extraneous enzymes — enzymes consisting of protein molecules, for example.

It was naturally a surprise, in the late 1970s, that it should have been discovered that genes of eukaryotic cells — stretches of double-stranded DNA helix — are not always connected pieces of helical DNA, and that functional pieces of DNA ("exons", where "ex-" stands for "expression") should be interrupted by apparently irrelevant stretches (called "intervening sequences", or IVSs).

Logic and cellular economy would argue that these irrelevancies should be irrelevant from the start, and thus never converted into RNA, but that is not what happens in real cells. Instead, the full length of a gene, introns as well as exons, is converted (transcribed) into a full-length molecule of RNA ("pre-mRNA") which is in turn converted into a molecule of mRNA somewhere in the nucleus. From the outset, this process (called splicing) has been a philosophical puzzle: for RNA transcripts to be accurately spliced, either the nucleus of a eukaryotic cell must contain a specification, distinct from the genetic DNA itself, of which stretches of RNA are to be excised, or the RNA transcripts themselves must contain a specification of how their intron sequences are to be removed.

In mammalian cells, many splicing processes turn on the recognition of splicing sites in RNA molecules by nuclear particles called small ribonuclear particles

(RNPs). But it is now also plain that, for at least some introns, self-specified self-splicing is the rule. In other words, some introns embody both their own specification and the molecular mechanism by which they are excised. Dr Thomas Cech and his associates, of the University of Colorado at Boulder, have contributed substantially to this discovery in a series of articles in the international literature beginning in 1982. They have worked out the details of a mechanism by which introns are excised from RNA transcripts by means of intramolecular reactions between the nucleotides of the introns and those of the adjoining exons.

Among other things, Cech showed that some RNA molecules, entirely innocent of protein from which all previously known enzymes are made, can function as enzymes. He coined the term "ribozyme" to describe such enzymes consisting solely of RNA. The interest and importance of the discovery was recognized by the Nobel committees this year, which have awarded Cech the chemistry prize. The award is widely applauded. One member of the Canberra group whose work has prompted the Australian patent application says "we were astonished" by Cech's first evidence that introns could be excised from RNA molecules without the intervention of protein enzymes. Cech's Nobel prize, the

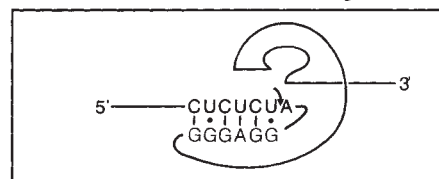


FIG. 1 Upstream splice recognition. The nucleotide sequence —CUCUCUA— is the end of the upstream exon, otherwise represented by the line labelled 5'. The nucleotide A is the first of the intron sequence, also represented as a line except for the six nucleotides —GGGAGG (reading from left to right), which are nucleotides 22 to 27 of the intron sequence in reverse order. Short vertical lines between pairs of nucleotides represent Watson-Crick (hydrogen) bonds between pairs of nucleotides. The directionality of polynucleotide molecules, whether RNA or DNA, is intrinsically determined by the way in which the linking phosphate groups can be attached either to the 5' or to the 3' position in the ribose (deoxyribose) ring of the sugar part of each nucleotide; because most, but not all, natural processes affecting polynucleotide molecules begin at the 5' end, that is also called the upstream end or direction. From the Cech, Zaug and Been patent application.

same researcher says, is "thoroughly deserved".

By any standards, the natural self-excision process is remarkable. It may be molecular biology's equivalent of Claude Shannon's famous trick box which appeared to be a closed cube about 30 cm on each edge presenting only a single toggle switch to the outside world. Toggling the switch would make the lid of the box open, when a robot arm would emerge, extend itself towards the switch so as to turn off the toggle and would then retreat whence it had come. Self-splicing introns similarly embody the means of their own quiescence.

For an intron to be capable of removing itself from a pre-mRNA molecule, three separate functions are required. First, it must be able to recognize its own beginning (or the end of the adjoining exon) and of cutting the molecule of which it is a part at exactly that point. The same function is required at the other end of the intron. And the intron must also be able to join together (by 'q ligation') the ends of the two exons which surround it. The first of these is illustrated in Fig. 1; the end of the adjoining exon is recognized by a short sequence of six nucleotides near the front of the intron, four of which, beginning at position 22 of the intron, form complementary pairs in the Watson-Crick sense. Given the importance of directionality (see Fig. caption) in the processing of all polynucleotide molecules, it is significant that the two short complementary RNA sequences that pair together have the opposite sense (as do, for example, the two strands of a DNA molecule).

Much of Cech's work has been concentrated on the mechanism of self-splicing of a particular intron in an RNA molecule destined to be a component of ribosomes (cellular organelles at which mRNA molecules direct the synthesis of the protein molecules that they correspond to) in the ciliate organism *Tetrahymena*

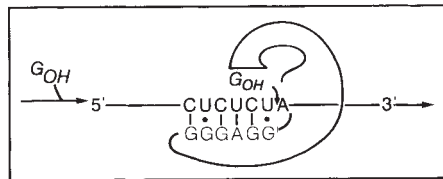


FIG. 2 A ribozyme for cleaving RNA. The curved line represents the ribozyme L-19 IVS, consisting of 395 nucleotides (414 minus 19). The internal recognition sequence pairs with a complementary sequence in the RNA molecule, represented by the straight line. After the reaction, the ribozyme molecule is left intact, the upstream section of the RNA molecule is released without a trailing phosphate group after the —CUCUCU sequence and the downstream section of the molecule has a guanine nucleotide added to the 5' end, said in the patent application to be a means for ready identification by means of its radioactivity, for example. From the Cech, Zaig and Been patent application.

thermophila. Cleavage from the upstream exon follows the process shown in Fig. 1. It is interesting that the downstream recognition process has only much more recently been described in detail (see Michel, Hanna, Green, Bartel & Szostak *Nature* **342**, 391; 1989).

In real life, the end result is that the intron, which is 441 nucleotides long, is excised from the pre-mRNA molecule, ending up as a circular RNA molecule incapable of interfering with the functioning of those other RNA molecules carrying nucleotide sequences recognized by the self-splicing intron. Such an intron, by definition, is not a true enzyme: it effects the transformation of only a single RNA molecule before lapsing into inactivity.

The crucial part of Cech's work, and that on which the US patent application is based, is his proof that parts of the self-splicing intron that do not lapse into the inactive circular form may still recognize particular nucleotide sequences in an RNA molecule, and then cut the molecule at that point. Because these molecular fragments survive the reaction intact, they are true enzymes, capable of catalysing several molecular reactions before being rendered inactive by extraneous causes. These are the materials to which the term "ribozyme" strictly applies. In principle, they are the equivalents in the chemistry of RNA molecules of the restriction enzymes (bacterial endonucleases) widely used in molecular biology for cutting DNA molecules at places of predetermined sequence.

By 1986, Cech and his associates had shown that it is possible to extract from their intron a piece (the whole intron less 19 nucleotides at the beginning) which is capable of carrying out the splicing of pre-mRNA molecules, normally an intramolecular reaction, by means of a more familiar intermolecular reaction (*Science* **231**, 470; 1986).

Applications

Very little imagination is required to conclude that ribozymes on this pattern could be used to manipulate the function of genes or even of entire cells. Ribozymes capable of cutting RNA molecules anywhere would be equivalent in their function to the protein enzymes known as ribonucleases, but ribozymes with built-in specificity for particular sequences of nucleotides along the length of RNA substrate molecules might be used to prevent the functioning of particular genes while leaving that of others unimpaired. One huge and important prize would be the use of small RNA molecules for interrupting the functioning of human retroviruses, that responsible for AIDS in particular.

It is not surprising that these discoveries should have provoked an interest in their commercial application. They are the

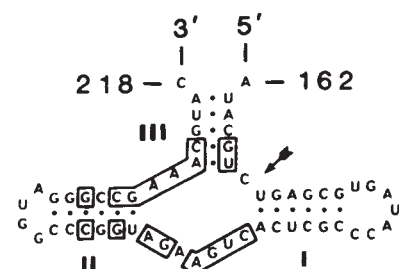


FIG. 3 Hammerhead self-splicing RNA. The hammerhead structure identified by Foster and Symons (see text) in the plus RNA strand of the virusoid VLSV, an RNA molecule of 324 nucleotides. The structure is inferred by allowing for Watson-Crick pairing between complementary nucleotides, but cleavage is found to occur uniquely between nucleotides 167 and 168 (marked with an arrow). Identical or similar secondary structures are found in other virusoids.

basis of the patent application first notified to the US Patent Office on 3 December 1986 by Thomas Cech and his two associates, Arthur J. Zaig and Michael D. Been. Under the procedures of the Patent Cooperation Treaty, University Patents Inc. also applied for patent protection in the chief countries of Western Europe and Japan, but not (apparently) in Australia; that application, filed internationally on 2 December 1987, was published (as WO 88/04300) by the Geneva office of the World Intellectual Property Organization on 17 June 1988 (PCT/US87/03161).

Two arcane features of the international patents regime complicate assessment of these applications. First, patent applications in the United States are not made public (nor is the fact of a rejected patent application disclosed); that means that there is no way of telling what protection has been sought in the United States, nor whether the international patent application now published differs substantially from any application that may have been made in the United States. US patent law also differs from that elsewhere in allowing patent applications to be lodged up to a year after the publication of relevant findings in the scientific literature, while most other patents authorities regard prior disclosure as a bar to the grant of a patent.

The international version of the Cech, Zaig and Been application has already been encumbered by references to papers published by Cech's own group in the twelve months before the original filing date in December 1986, which may or may not be obstacles to patent protection outside the United States. Uncertainty on such questions seems commonplace in the patents business. Most patent applications make a long series of claims (the Cech application lists 73) ranging hierarchically from the general to the particular, only some of which may be anticipated by particular publications, while it is always possible that applicants may persuade

patents examiners that their provisional judgements are mistaken.

The drift of the patent application rests squarely on the discovery, reported in January 1986, that the intron in the *Tetrahymena* pre-mRNA molecule can catalyse intermolecular reactions, and in particular carry out all the processes required for the removal of the intron, of which the most striking is the process by which an RNA molecule is cleaved at a predetermined sequence of nucleotides. The active ribozyme is the natural intron from whose 5' end 19 nucleotides have been removed, and which is called L-19 IVS — L (for "linear"), "-19" (meaning "minus 19") and IVS (for "intervening sequence"). Figure 2 shows the functioning of this ribozyme in what may be its chief role as an RNA-restriction enzyme, or endonuclease.

The experimental conditions are important. Ions of Mg^{2+} are necessary, as is either guanosine (abbreviated G) or its phosphorylated derivative GTP. What then happens in self-excision is that, by a mechanism not understood, the RNA chain is broken after the recognition sequence of six nucleotides at the downstream end of the preceding exon.

Although the Cech, Zaugg and Been patent application seeks protection for the enzyme functions of L-19 IVS other than its capacity to cleave RNA molecules — that of removing phosphate groups from the 3' ends of RNA molecules, for

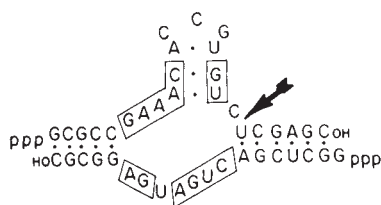


FIG. 4 Uhlenbeck's ribozyme. Two RNA strands synthesized by analogy with ASBV viroid sequences involved in a hammerhead self-cleaving region shown in the configuration in which they associate by means of Watson-Crick pairing. The lower strand is the 19-nucleotide ribozyme referred to in the text.

example — they are less relevant to the Australian patent application and likely to be, in any case, less important. In the list of 73 claims in the Cech, Zaugg and Been patent application, the second seeks protection for an "RNA enzyme" with "sequence-specific endonuclease" activity. Other more specific claims would protect the use of modified versions of the *Tetrahymena* ribozyme: the first two and the last of the 395 nucleotides may be eliminated from L-19 IVS without impairing its cleavage activity, while several examples are given of how the active (recognition) site of the enzyme can be modified so that it will cleave RNA molecules at places other than the join between a natural exon and an intron.

One question certain to arise in the

months ahead is whether the claim for protection of endonuclease activity in the Cech, Zaugg and Been application is so general that all further proposals for the use of ribozymes will be somehow subsidiary, at least in the United States (where Cech's prior publications are not, on the face of things, a bar to patent protection). The issue is not simply a matter of patent law and procedure; much will also depend on the commercial judgements of potential challengers of the relative costs of a challenge and a commercial compromise.

Further development

Apart from prompting a patent application, Cech's discoveries in the period 1982–86 seem powerfully to have stimulated other research groups already then concerned with the mechanism of self-splicing. One of these is the group of Robert H. Symons at the University of Adelaide, which has a long-standing interest in the mechanisms of splicing in RNA molecules, especially the RNA plant viruses and the RNA parasites of plants known as viroids and virusoids. Haseloff, one of the authors of the CSIRO patent application, was a member of this group as recently as 1982.

Symons and his associates appear to have succeeded in defining the necessary and sufficient structural requirements for self-cleavage in several of these plant RNA molecules. The results are neatly illustrated by the self-splicing of the virusoid of the lucerne transient streak virus (LTSV), which in its intact form consists of several virusoid particles, each a single-stranded RNA molecule 324 nucleotides long packaged, together with an RNA molecule ten times as long, into an icosahedral structure. The virusoids require the presence of the longer RNA molecule of the LTSV, and may themselves be formed as inactive rod-like particles of RNA in which the infective virusoid molecule is bound for most of its length to a Watson-Crick complementary "minus" strand. This rod-like structure seems to be capable of splitting into two separate strands by self-cleavage (see Anthony C. Foster and R.H. Symons, *Cell* 49, 211; 1987).

What Symons and his associates have shown is that both the plus and minus strands of the virusoid RNA (called vLTSV) are also capable of self-cleavage at specific places along their length, and that the sequence of nucleotides around that point will allow for a distinctive secondary structure of the RNA molecule, called graphically a "hammerhead" structure (see Fig. 3). Similar structures have been recognized in other plant virusoids, leading Symons and his colleagues to conclude that it is a structure prone to self-cleavage in which Mg^{2+} ions assist (but are not always necessary).

There is no suggestion that these hammerhead structures, recognized in

RNA molecular parasites of plants, are involved in all RNA self-splicing; Cech's L-19 IVS is one exception; it is merely that the hammerhead structure seems to be a sufficient condition for self-splicing in several natural RNA molecules. The work of Symons and his colleagues has been concerned exclusively with naturally occurring processes, and not with the design of molecules of RNA capable of acting independently as ribozymes in the strict sense.

Observations such as these show that a quite small stretch of RNA may be sufficient for self-cleavage, and it is natural that they should have helped to stimulate an attempt to define the minimum self-cleaving region. One important and relevant attempt in this direction is that of Ohlke C. Uhlenbeck from the University of Colorado at Boulder, a colleague of Cech's but not a member of his research group (see *Nature* 328, 596; 13 August 1987).

Noting the demonstration by Symons and his colleagues of the effectiveness of hammerhead structures in the cleavage of plant parasite RNA, Uhlenbeck synthesized two RNA molecules with nucleotide sequences corresponding to those constituting the hammerhead structure in avocado sunblotch viroid (ASBV), showing that the shorter RNA molecule consisting of only 19 nucleotides would cleave its longer companion at the place predicted by what happens in ASBV itself (see Fig. 4).

That neat demonstration is likely to play an important part in the resolution of the status of the US and Australian patent applications, as the news report of the 12 October explained. Uhlenbeck's 19-nucleotide RNA is a true ribozyme, catalysing the cleavage of more than one molecule of its substrate. On the face of things, it appears also to be a particular example of the class of ribozymes at the core of the Australian patent application, which had not then been lodged. Although Uhlenbeck has not applied for a patent, and does not intend to do so (even under US law, he would be out of time), the publication of his paper may be held to invalidate later applications for patents covering ribozymes of similar design, or even to show that developments such as these were, by mid-1987, "obvious" in the sense in which patent law uses that term.

CSIRO's application

The Australian patent application, first lodged with the Australian Patent Office on 15 December 1987, is a model of clarity in such complicated circumstances. Observing that self-cleavage sites have been recognized in several plant viruses and related RNA, the document goes on to suggest a means by which efficient ribozymes may be constructed from synthetic RNA by an elaboration of the hammer-

head principle.

Indeed, its authors generalize what is known of the common nucleotide structure of self-cleaving RNA molecules. Their claim rests on the assertion that ribozymes of any pre-determined specificity may be constructed by synthesizing RNA molecules in which the active cleavage element of the ribozyme is surrounded by short stretches of RNA capable of attaching, by Watson-Crick complementarity, to RNA sequences on either side of the desired cleavage site in the substrate RNA molecule (see Fig. 5). The essential elements of the patent application have been published (W. L. Gerlach and J. P. Haseloff, *Nature*).

Significantly, the Australian patent application (PCT/AU88/00478, published as WO89/05852) acknowledges Cech's earlier work and the existence of his patent application, but asserts that the two sets of ribozymes are distinct. First, the Australian application says, Cech's ribozyme occurs naturally in *Tetrahymena thermophila*. But the argument continues that Cech's ribozymes require the presence of guanine or one of its derivatives (usually GTP), add this nucleotide to the upstream end of the downstream cleaved fragment and holds that the Cech ribozyme is likely to be less specific than the Australian because fewer nucleotides can be involved in the hybridization of the ribozyme and its substrate molecule, especially because some of the crucial nucleotides cannot apparently be changed. The result, the Australian application says, is further to restrict the freedom to engineer specificity.

The Australian patent application is also more specific about the applications foreseen for artificial ribozymes than are Cech, Zaug and Been. In that document, the only use of ribozymes explicitly described is as laboratory reagents. But the Australian patent application says that ribozymes can be engineered to cleave any RNA molecule at a pre-determined sequence of ribonucleotides, meaning that target RNA sequences can be made inactive. Alternatively, as the case is put, ribozymes might be produced endogenously in cells by the transfer of self-replicating plasmids or by microinjection.

The application says explicitly that infectious mammalian viruses might be inactivated by the direct administration of appropriate ribozymes or by incorporating into vectors such as vaccinia virus carrying sequences of DNA, that would allow the production of the ribozymes endogenously. The same (the application says) could be done to neutralize the effects of unwanted gene activity in plants, even to the extent of engineering plants that produce fruit without stones, as well as to treat genetic disease in people where the underlying defect is the overproduction of a protein produced at the prompt-

ing of some mRNA.

Nobody can cavil at the importance of these claims. If CSIRO and its commercial partners — CSIRO hopes there will be an Australian partner as well as the French company Groupe Limagrain — can find a way of delivering specific ribozymes into malfunctioning cells, they will have found a way of curing many of the world's ills. There may be practical difficulties to be overcome before engineered ribozymes can be used as medicines, and regulatory hurdles to be surmounted before vaccinia virus and other genetic vectors can be used in the prophylaxis of human disease, but CSIRO's enthusiasm for its "gene shears" process should not be crabbed.

Disputations

Nobody, unfortunately, can tell in advance how patent lawyers and (if it comes to that) the courts will determine the relative merits of these rival, but not necessarily conflicting, claims. Cech, Zaug and Been were first, by more than two years, but Gerlach, Haseloff and Jennings describe what appears to be a more efficient ribozyme, and one whose specificity may be pre-determined by a knowledge of the nucleotide sequence of the RNA it is meant to cleave, irrespective of its source.

Variations of patent law from one country to another will complicate the resolution of the issue; for example, patents might be awarded in the United States to both applicants, but to one or other elsewhere. On the face of things, the existence of independent prior publications, such as those of the Symons group and particularly of Uhlenbeck, would seem to be an embarrassment for CSIRO's case. On one view, the commercial value of CSIRO's gene shears technique may hang on the speed with which the organization is able to work towards its practical exploitation.

How did *Nature* come to be calling the odds on such a complicated issue in its news report of 12 October? It made mistakes, listed in what follows, in incompletely extenuating circumstances.

One error was that of CSIRO, which first announced at the end of July this year that it had signed an agreement with Groupe Limagrain for the exploitation of its "patent" on gene shears. If that claim had been less definite (only an application for a patent had then been lodged), *Nature's* informants in the United States might have been less eager to say that something must be wrong. The outcome was a conversation with the Boston-based patent attorneys acting for US Biochemicals Corporation, which has acquired rights to exploit the Cech, Zaug and Been patent.

Six weeks went by before *Nature's* Australia correspondent, Tania Ewing, who had by then spoken to the US patent

attorneys, approached Dr Jim Peacock, at the Plant Industries Division of CSIRO, telling him at the outset only that doubts had been cast on CSIRO's announcement in July, but in later conversations asking for his reaction to the US claims that the Cech patent application had precedence over whatever might be claimed in Australia. The response appears to have been unspecific; detailed discussion was declined, but Dr Peacock and other members of the gene shears team insisted that they were "confident" that their patent application would succeed, pleading CSIRO's patent lawyers in their support. But access to the lawyers was denied, on the grounds that "that is not CSIRO policy".

The contemporary sociology of Australia is relevant. The country is lavishly over-governed, but the press is constitutionally powerful; public officials who refuse to answer questions on the telephone may be lampooned. Meanwhile, CSIRO has made a painful transition from semi-autonomous research organization to government agency of a kind; the pressure to succeed commercially is intense and insistent, as is the fear of failing to succeed. So researchers dutifully answer the telephone, even returning calls from journalists, but deal cautiously with questions of a kind that politicians might understand.

If Tania Ewing had been offered a sight of the two disputed patent applications, she would not have concluded that her US patent attorney informants were correct in their assertion that the Cech patent would supervene and that CSIRO's confidence was misplaced; instead, she would probably have decided that there is a tangled fight ahead. Yet even two weeks ago, CSIRO was still saying that it would be possible to secure a copy of the Cech patent application only from the International Patent Office at Geneva, its own complement of photocopying machines notwithstanding.

Those, so far as *Nature* is concerned, are most of the extenuating circumstances. Three errors in the 12 October account stand out, as follows:

■ Dr Jim Peacock was at one stage asked whether he "remembered a letter from a man called Jim Chen", allegedly a discouraging report on the Australian patent application from the Australian Patent Office. But in reality, the document concerned was not a letter but a standard form report (citing the Uhlenbeck research as potentially prior art), with Jim Chen's signature in the bottom right-hand corner. Peacock had indeed forgotten (if he had ever remembered) that the form had been signed by Chen. When reported to have "forgotten," the results of this international patent search, he was made to look a fool. He was, of course, fully aware that he had received the research report.

■ In the same vein, Mr Stephen Castle at the Australian Patent Office was quoted directly as saying "We don't think they [CSIRO] will get their patent". There is a direct conflict of evidence on this point between what Mr Castle says occurred and Tania Ewing's notebook. Mr Castle says that on the day Ms Ewing reached him by telephone, he was standing in for his immediate supervisor, Jim Chen. The question was whether the "X" rating in respect of Uhlenbeck's paper would be fatal for CSIRO's application. Mr Chen says that he explained he had no direct knowledge of the application, that he explained that he could not discuss particular applications, but recalls saying that such an annotation "means that a patent will probably not be granted". At an interview conducted in the presence of Mr Murray Heddrich, Deputy Commissioner of the Patents Office, Mr Castle also says that he explained that an "X" document would invalidate only the particular claims to which the annotation referred (a third of the total in the case of the Australian patent application) and that applicants might in any case successfully challenge the examiner's judgement.

There is no simple way in which such a conflict between written and remembered evidence can be reconciled. But it can be understood. In Ms. Ewing's notebook, the quotation "We don't think they'll get their patent" follows directly after a notation of a question about Uhlenbeck's article, implying that the two questions were closely linked. Mr Castle's recollection is that they were widely separated. Whatever the intervening material, it goes without saying that anybody hearing the phrase "We don't think they'll get their patent", even if in response to a question about the general significance of "X" ratings, would have concluded that what applies to the general also applies to the particular, only recently discussed. But Mr Castle's reservations, however emphasised during that conversation still apply: prior publications threaten only the claims to which they are judged relevant, and that judgement may always be challenged. The moral seems to be that, when writing about patent applications, it is important to read them (if you can get hold of them).

■ In the United States and Western Europe, *Nature* distributes each week to interested journalists a summary of each week's scientific content, regretting that the timing of this operation is such that important items of news cannot be thus advertised. On this occasion, it was agreed between Sydney and Washington that advance notice of the news report of the gene shears patent should be distributed to a few Australian journalists, which had the unfortunate effect of amplifying the misleading aspects of the news report.

The first of these issues is a misunder-

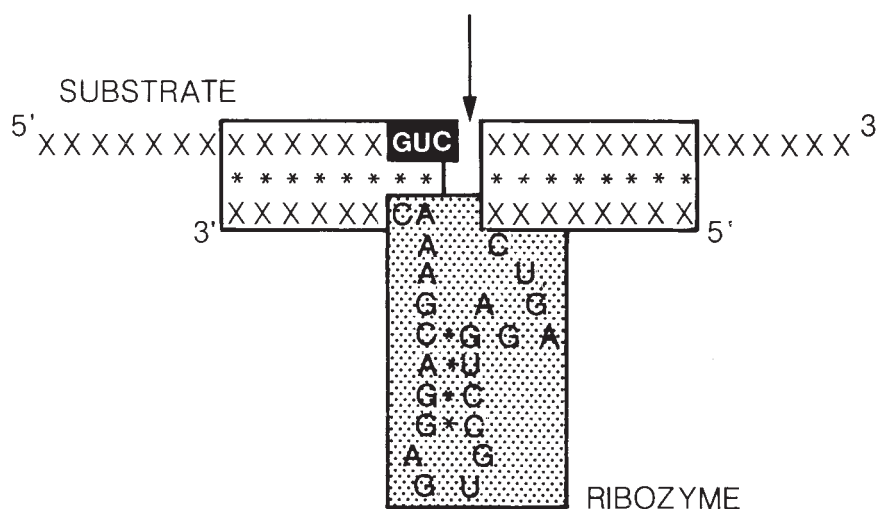


FIG. 5 The CSIRO ribozyme. Schematic diagram (from the Australian patent application) in which X stands for any nucleotide and X.X for any (or either) pair of complementary nucleotides. The shaded block includes the shortest RNA sequence found to be effective in cleaving RNA parasitic on plant cells. As shown, the ribozyme would cut the upper substrate RNA strand after the nucleotide triplet GUC, but the patent application argues that this restriction may be relaxed so long as the central nucleotide in the triplet remains as uracil and so long as the third is not guanine. The RNA sequences attached to the common ribozymal unit would be synthesized so as to be complementary in sequence to the naturally occurring substrate and thereby enhance the specificity of the ribozyme.

standing, the second a potential misunderstanding, the third is an error of judgement. *Nature* is sorry for the hurt they may have caused, however trying the circumstances. But it is hard to fathom the ferocity of the Australian reaction. Mr Barry Jones, the federal Minister of Science, telephoned Ms Ewing on the afternoon of 19 October to say that the Australian government would sue her and *Nature* for what she had written, which information was the basis for at least two daily newspaper reports the following day. Two weeks ago, Mr Jones rejected the charge that he was seeking to intimidate journalists, but could not recall the circumstances in which he had made the call.

Nature has the fullest confidence in its Australia correspondent and hopes to be able to do so for many years. There are nevertheless several lessons to be learned from this saga, as much by people in offices in Washington and London as by reporters in the field. One is the old journalistic saw that one cannot be too careful — of volunteered information (although, without it, journalism would grind to a halt), of the contents of documents which have not been read, but only reported by potentially interested parties, and of the difficulties of pursuing information on two continents which are separated by 16 hours on the clock. On this occasion, proper scepticism of Australian claims seems to have been unfairly magnified.

The plain truth is that the Gerlach and Haseloff design for ribozymes is certain to be widely applied in the laboratory processing of RNA, while the potential of the applications in medicine and agriculture is potentially very great. Before these prizes can be won, of course, several less interesting practical problems will have to be

solved. Using ribozymes as medicines for the treatment of AIDS, for example, will require that means of delivering effective doses of the material to infected lymphocytes will have to be found. By all accounts, the pharmaceutical companies are not amused.

Yet Peacock says that the collaboration with Groupe Limagrain has already embarked on two development projects with promise. It would be good to think that Australia will yet win a good share of the benefits seemingly in sight. The pressure exerted by the federal government in such directions may often be over-onerous or even short-sighted, but it has been Australia's dream for several decades to supplement its staple exports of primary products from mines and farms with exports whose added value is greater. That, of course, is why neat gene shears were so generally acclaimed in July this year.

The question whether interesting science is now too much confused with the patents business is a perennial question but also often an academic one. For many people there is often no choice. Protecting intellectual property has become part of research, even survival in research. Yet the task of briefing patent attorneys is an irksome and time-consuming obligation from which some people still shrink. A predictable but unexpected consequence is the effect of the patents business on people's demeanour. Much of the reticence encountered in describing CSIRO's gene shears seems to have arisen both from unfamiliarity with arcane patents practices and from the way in which attorneys frequently swear their clients to silence. Too much of that could bring great trouble. □

Magma transport through dykes

Greg A. Valentine

THE transport of magma through the Earth's crust, whether from relatively shallow magma reservoirs (depths of a few kilometres) or the upper mantle (depths approaching 100 km) has always presented geologists with a challenging problem. Much of the challenge arises from the complex behaviour of magmas under conditions of changing temperature, pressure and composition. An important aspect of magma flow through fissures is the exchange of heat with the surrounding rocks. For flow to be sustained, the heat advected from the magma reservoir must successfully compete with the loss of heat to the cooler country rocks. This competition leads to a physical system in which hydrodynamics (the

is within dykes. Oceanic islands within tectonic plates, such as Hawaii, are characterized by large, low-relief shield volcanoes that erupt magmas from linear fissures on their summits and flanks (see figure); these fissures are the surface expression of dykes. Even silicic magmas are commonly influenced by flow in dykes at some level in the crust. Examples of this include dykes that radiate out from central conduits of stratovolcanoes, linear systems of silicic domes and ring dykes associated with large intrusions.

Magma transport via dykes presents two problems which can, to some extent, be treated separately. The first is formation of the crack. In many cases the initial crack may be a pre-existing fracture in the

given composition, the viscosity depends on the temperature and strain rate. As a magma cools or heats, it can undergo drastic phase changes that modify the behaviour of the fluid and that involve significant exchanges of latent heat. Previous treatments of magma cooling due to heat loss to the dyke walls have consisted of simple static conduction calculations³, empirical calculations of heat exchange between flowing fluid and a wall for one-dimensional flow⁴, and two-dimensional calculations allowing for some advective effects⁵. None of these formally allowed for the effects of latent heat released or absorbed by material as it solidifies or melts, respectively.

Bruce and Huppert¹ make an important step in the theoretical treatment of magma flow in dykes by allowing the advection of heat by flowing magma to affect the temperature distribution in the wall rock. They include melting, solidification and

related latent-heat effects in their analysis.

Two important regimes are found: one in which heat loss to the walls exceeds heat advection and the dyke freezes shut; and a second in which initial heat loss may exceed advection, but in which advection and latent heat effects eventually win and the dyke remains open and may even grow wider. This second regime involves initial solidification of magma along the walls followed by re-melting (meltback).

Bruce and Huppert's calculations have a wealth of implications for eruption dynamics

A typical basaltic fissure eruption, on Mauna Ulu, Hawaii, 1969, showing fire-fountaining due to eruption of gas-rich magma. These eruptions commonly occur along an entire fissure during initial phases, but eventually become localized to a few discrete vents along the fissure. According to Bruce and Huppert¹ this trend may be the result of extreme sensitivity of magma flow in fissures to the initial width of the fissure. Small irregularities in the

initial width may become pronounced as an eruption progresses, leading to concentration of magma discharge in slightly wider parts of the fissure. (Photograph by W.A. Duffield, US Geological Survey.)

advection of heat) and thermodynamics (latent heat from solidification and diffusion of heat into surrounding rocks) are strongly coupled and cannot be treated as independent processes. In an important new analysis, Bruce and Huppert show on page 665 of this issue¹ how the coupling can lead to complex processes of solidification followed by re-melting. They derived a criterion that separates dyke blockage by solidification from sustained flow.

Field evidence shows that magma is transported through the crust mainly by the flow of magma through cracks, or dykes. Hence, dykes are a significant feature in nearly all magmatic provinces of all ages, especially where mafic or ultramafic (iron- and magnesium-rich) magmas are present. In areas undergoing tectonic extension, such as mid-ocean ridges and continental rifts, the dominant means of magma transport to the surface

host rock, but in others the fracture is probably propagated by the magma itself in a manner similar to hydrofracture processes. This magma-fracture process has been studied by Spence and Turcotte² for highly simplified cases that can be mathematically analysed by a similarity solution. They find that magma can propagate fractures in a vertical direction at speeds of the order of $0.01-100 \text{ m s}^{-1}$, the buoyancy of the magma driving them upward. Their conclusion is that most of the transport problem resides in the flow of magma once the fracture has opened; the initial fracturing itself is not, in general, a limiting process.

The fluid-flow problem is more daunting owing to the nature of magmas. The rheology of magma is strongly dependent upon its composition. Viscosity, for example, can vary over several orders of magnitude for a change in SiO_2 content of just 20 per cent. Furthermore, at any

and petrological variations of eruptions from fissures. For example, the sequence of initial solidification followed by meltback could strongly influence eruption dynamics. The solidification would concentrate volatiles into the remaining melt in a dyke, which upon eruption may appear to be 'artificially' volatile rich. In a basaltic fire-fountain eruption, this partitioning of volatiles may lead to a higher fountain, which would in turn lead to overestimates of the volatile content in the magma chamber. Later, as meltback occurs, volatile-depleted material (the old chilled margin) would be mixed up with the flowing magma to deplete the eruptive mixture 'artificially'. Such events could produce apparent variations in original magma composition, even if the original composition was unzoned. Is it possible that this process is responsible for fluctuations in fountain heights during eruptions such as observed by Head and Wilson⁶?

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

These processes could also influence the major- and trace-element composition of erupting magmas, which should be altered to some degree by the solidification-meltback process, especially if meltback reaches past the original dyke walls and involves original country rock. Also, they are likely to produce nonequilibrium textures in phenocrysts.

It seems essential to collect samples of basalts as they erupt from a given vent, in order to observe some of the processes that Bruce and Huppert predict, as cooled dykes are highly unlikely to preserve these processes. A difficulty arises because the original compositional zonation in the magma reservoir and the effects of draw-up into the dyke would have to be established independently of the basalt sampling — a challenging prospect.

Finally, it is worth noting that even though Bruce and Huppert's analysis is an important step in elucidating magma transport in dykes, we are still a long way from a physically complete understanding of the process. For example, none of the analyses mentioned above, including

Bruce and Huppert's, takes into account the real rheology of magmas. In addition, multiphase flows that result from the release of dissolved gas from the magma need to be incorporated into the heat-transfer problem. Neglect of these processes is attributable partly to poor data on magma properties, but also to limitations in mathematical approaches that rely mainly on analytical techniques. Numerical solutions of the full set of governing equations with minimal simplifications would constitute a significant advance. □

Greg A. Valentine is in the Geoanalysis Group, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA.

1. Bruce, P.M. & Huppert, H.E. *Nature* **342**, 665–667 (1989).
2. Spence, D.A. & Turcotte, D.L. *J. geophys. Res.* **90**, 575–580 (1985).
3. Wilson, L. & Head, J.W. III *J. geophys. Res.* **86**, 2971–3001 (1981).
4. Spera, F.J. *Contrib. Mineral. Petrol.* **88**, 217–232 (1984).
5. Delaney, P.T. & Pollard, D.D. *Am. J. Sci.* **282**, 856–885 (1982).
6. Head, J.W. III & Wilson, L. *J. geophys. Res.* **92**, 13715–13719 (1987).
7. Vergiollie, S. & Jaupart, C. *J. geophys. Res.* **91**, 12842–12860 (1986).

CELL BIOLOGY

The pushy ways of a parasite

John E. Donelson and Alice B. Fulton

A SERIES of electron micrographs in the *Journal of Cell Biology*¹ provides new insight into how an intracellular bacterial parasite, *Listeria monocytogenes*, behaves inside cultured macrophages and how it moves from one macrophage to another without ever leaving the cytoplasm. The authors of the paper, Tilney and Portnoy, not only demonstrate the cunning way in which the parasite traverses the perilous divide between host cells, but provide pointers to fresh approaches for studying the cytoskeleton.

All parasites of mammals must evade the immune systems of their hosts. Some do it by hiding within host cells, but this strategy has its own complications as the insides of host cells can be distinctly hostile². Parasites that survive within host cells such as macrophages have therefore evolved a variety of mechanisms to prevent themselves from being destroyed by cellular defences such as lysosomal acid hydrolases and the oxidative response. For example, some intracellular parasites inhibit fusion between the parasite-containing phagosome and lysosomes (*Toxoplasma*, *Chlamydia*); others are resistant to lysis by lysosomal enzymes (*Leishmania*, mycobacteria); and still others escape from the acidic 'phagolysosome' into the host cell's cytoplasm (*Listeria*, *Trypanosoma cruzi*)³. But even after coping with the host cell's defences, the parasite must still avoid destruction by the humoral immune system during transfer from one

cell to another.

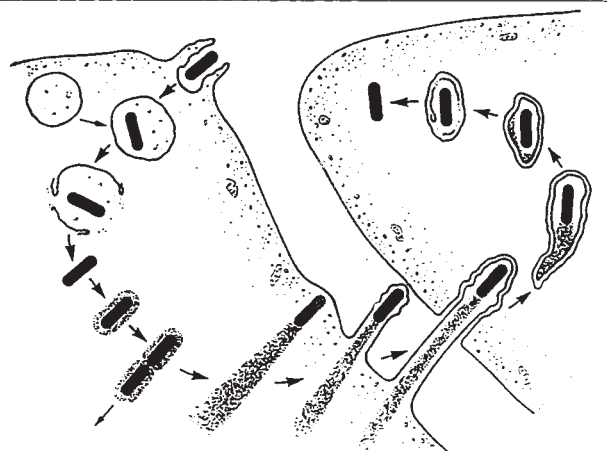
Listeria monocytogenes is a gram-positive, non-spore-forming pathogen of humans, sometimes found in unpasteurized milk, that penetrates macrophages, fibroblasts and enterocytes. Following phagocytosis, the parasite-containing endosome fuses with one or more lysosomes (see figure). Tilney and Portnoy's electron micrographs show that the phagolysosome membrane dissolves before the bacterium is damaged; the membrane is probably degraded by a bacteria-encoded haemolysin, because *Listeria* mutants lacking haemolysin fail to escape from the phagolysosome⁴.

After phagolysosome breakdown, the *Listeria* subverts the macrophage's cytoskeleton for its own purposes in an unexpected way. First, a halo of actin filaments (which are decorated by myosin subfragment-1) surrounds the bacterium. After the *Listeria* divides, this halo becomes asymmetric, forming a tail at one end of the bacterium and taking on the appearance of a short comet. The halo and comet tail are densely packed with short randomly ordered filaments. As the tail elongates, the bacterium is apparently 'pushed' by the comet tail towards the macrophage surface. Once at the surface, the *Listeria* distends the host membrane to form a pseudopod-like extension that touches another macrophage. The *Listeria* and some of the actin tail are engulfed by the new host cell and another round of infection begins. Thus, once safely inside the first macrophage, the *Listeria* need not confront the perils of the humoral immune system again.

This exploitation of the host's cytoskeleton raises several questions. First, how does the *Listeria* induce the macrophage's actin to polymerize? The halo and comet tail are not simply pre-existing filaments recruited into these structures, because cytochalasin D prevents their formation and stops the bacterium from moving towards the surface. The filaments appear to grow from some component of extracellular material around the *Listeria* as all the filaments begin near the bacterial surface. *Listeria* has recently been found to export a basic (isoelectric point 9.3) protein of relative molecular mass 60,000 involved in bacterial attachment to the host cell⁵; perhaps this positively charged protein also nucleates actin polymerization in the host cytoplasm, as does polylysine *in vitro*⁶.

Second, why are these particular filaments so short, skewed and stable? Their average length appears to be only a few tenths of a micrometre, yet the comet tail of several micrometres is stable for up to an hour. Does *Listeria* secrete a toxin, such as

Steps in the entry, multiplication, movement and transfer of *Listeria* from one macrophage to another. Once inside a macrophage the bacterium escapes into the host's cytoplasm and is surrounded by a halo of the host's actin. After division, the actin around the daughter bacterium becomes asymmetric, taking on the appearance of a comet. As the tail elongates, the bacterium is apparently propelled towards the macrophage membrane, where a pseudopod-like extension of the cell membrane is formed. When the pseudopod is engulfed by another macrophage, the bacterium is released into its new host cell.



phalloidin, that stabilizes filaments? If so, how does the rest of the actin in the cell escape its effects? Why is the long comet tail composed of these short, skewed filaments when normal elongated actin structures generally have longitudinal order, sometimes approaching a paracrystalline state? Does *Listeria* secrete a bundling protein to hold together the short, skewed filaments, or does it recruit one from the host?

Perhaps the simplest explanation is that *Listeria* has many high-affinity nucleation sites for actin polymerization on its surface. The progeny bacterium appears to have fewer sites at one end so that actin polymerization imparts motion to the bacterium by mechanically pushing it. Actin polymerization is known to be involved in pushing out the membrane of the acrosomal process of *Thyone briareus* sperm⁷.

Regardless of the mechanism underlying the bacterium's movement, how it becomes directed towards the macrophage surface remains unexplained. However, if the nucleating component is shed by the bacterium as it moves, the filaments would be left behind in the tail. The apparent stability of the filaments may be a consequence of their exceedingly high density and the presence of the nucleating

PALAEOCLIMATE

component. Their disarray may reflect the absence of crosslinking proteins.

At least some of these questions can be addressed at the genetic level because transposon mutagenesis and plasmid DNA transformation can be carried out with *Listeria*⁸. Furthermore, cytoskeletal subversion may be a common tactic used by other intracellular parasites. Intracellular infections by bacterial and protozoan parasites have long been used as models for studying cell-mediated immunity — Tilney and Portnoy's findings indicate that they are also good systems for examining cytoskeletal formation, organization and function. □

John E. Donelson of the Howard Hughes Medical Institute and Alice B. Fulton are in the Department of Biochemistry, University of Iowa, Iowa City, Iowa 52242, USA.

1. Tilney, L.G. & Portnoy, D.A. *J. Cell Biol.* **109**, 1597–1608 (1989).
2. Moulder, J.W. *Microbiol. Rev.* **49**, 298–337 (1985).
3. Sher, A. & Colley, D.G. in *Fundamental Immunology*, 2nd edn (ed. Paul, W.E.) 957–983 (Raven, New York, 1989).
4. Portnoy, D.A., Jacks, P.S. & Hinrichs, D.J. *J. exp. Med.* **167**, 1459–1471 (1988).
5. Kuhn, M. & Goebel, W. *Infect. Immun.* **57**, 55–61 (1989).
6. Lin, D.C., Tobin, K.D., Grumet, M. & Lin, S. *J. Cell Biol.* **84**, 455–460 (1980).
7. Tilney, L.G., Bonder, E.M., Coluccio, L.M. & Mooseker, M.S. *J. Cell Biol.* **97**, 112–124 (1983).
8. Gaillard, J.L., Berche, P. & Sansonetti, P. *Infect. Immun.* **52**, 50–55 (1986).

Deep trouble for climate change

N.J. Shackleton

THE search for geological evidence of past climate changes has yielded much material for modellers trying to understand the principal factors controlling these changes, be they variations in the ocean–atmosphere system or slow changes in the Earth's orbit as suggested by Milankovitch. New data derived from the ages of fossil corals, presented by Fairbanks on page 637 of this issue¹, show that we still do not understand the very rapid change in climate during the period known as the Younger Dryas at the end of the last glaciation.

The fossil coral terraces of tropical islands provided the first convincing geochronological evidence that the Milankovitch hypothesis of long-period climate forcing might be correct². The uplifted coral tracts have yielded extremely valuable data on the history of climate-induced changes in sea level^{3,4}. Coral deposited over the past few thousand years has also been dated for reconstructions of the approach of sea level to its present position as the ice sheets of the last glacial maximum finally faded away⁵. By drilling into the uplifted interglacial terraces in Barbados, Fairbanks and Matthews⁶ were able to document on the ground the relationship between the

changing isotope composition of the oceans as ice sheets melted and grew, and the height of the sea relative to Barbados.

Despite these contributions, there has been a fundamental difficulty in interpreting the ocean oxygen-isotope record — a key to so many aspects of palaeoclimatology — because we have had no means of accurately comparing ¹⁸O and ice volume, or sea level, over the whole glacial–interglacial range⁷. Armed with his experience of the uplifted coral terraces, Fairbanks has now filled in much of this gap in our understanding by taking a look at the submerged coral, which one had always suspected must be there awaiting attention. In this issue¹ he shows the first, exciting results of that enterprise.

What Fairbanks has done is to drill and core submerged coral at depths below sea level ranging from 10 m to about 130 m. Concentrating on *Acropora palmata*, which grows only within about 5 m of the sea surface, he has obtained the ¹⁴C age and present depth of each piece. Correcting this depth for the slight uplift of the island over the past 15,000 years he obtains the first almost continuous record of sea-level rise during deglaciation (even a single well-dated estimate of the glacial low-stand would have justified publication).

Still more exciting is the fact that Fairbanks's temporal cover is good enough for him to differentiate the sea-level curve to present a direct measure of the rate of rise of sea level, which is of course equivalent to a measure of the meltwater flux into the oceans. As Fairbanks points out, this will be subject to some subtle modification when we have a more accurate knowledge of the relation between so-called 'radiocarbon years' and true years. However, his conclusions, even as they stand, call into question a recent interpretation of the Younger Dryas event.

The Younger Dryas event was an interval lasting a few centuries between 11,000 and 10,000 years before present (BP) during which the climate of northwest Europe, which had been quickly improving after the glacial maximum, was suddenly thrown back to near-glacial levels. Glaciers advanced again, an event marked as the Loch Lomond Stadial in Britain. The reason for this event has intrigued people for many years; one reason for devoting special attention to it is that whereas the growth of the great ice sheets took place over thousands of years, this event appears to testify to the ability of the global climate system to shift from an interglacial to a glacial state in a matter of decades. Is global climate balanced so finely that it could flip easily into a glacial mode?

In a series of papers^{8,9}, Broecker has argued that the Younger Dryas may have been a feature peculiar to the deglaciation period rather than something that could happen at any time. He recently demonstrated⁸ that the Younger Dryas occurred between two periods during which glacial meltwater flooded down the Mississippi river to the Gulf of Mexico, as indicated in cores recovered from the Orca Basin. Broecker's argument is that, between these two events, meltwater was diverted and instead entered the ocean so far north that it upset the density stratification of the North Atlantic.

His reasoning is that such a huge meltwater influx down the Saint Lawrence Seaway would significantly lower the salinity of the surface of the whole northern part of the North Atlantic. This would prevent the winter cooling of the sea surface in the Norwegian and Labrador Seas from raising the density of the water enough that it could sink to form the North Atlantic Deep Water. Broecker's underlying hypothesis⁹ is that the global climate system has two stable states, a glacial state and an interglacial state, and that upsetting deep-water formation in the North Atlantic is a mechanism for flipping the system into the glacial state. In this picture, the Younger Dryas would be such a flip into the glacial state that arose as a chance by-product of the geographically complex melting history of the North American ice sheet.

Fairbanks's work shows that, in fact, melting almost stopped during the Younger Dryas; the meltwater flux fell to a fifth of the maximum value it had had a millennium previously. The peak flux was at 12,000 years BP, identical in age with the greatest Mississippi flow⁸. By 11,000 years BP, the rate of sea-level rise was at a minimum, the same date as the cessation of enhanced Mississippi flow. Fairbanks's data show that the pulse of very rapid melting came to an end shortly after 12,000 BP; it would be very difficult now to argue that the diversion of the Mississippi flow could have been the main trigger for this important climatic episode, and it seems likely that the whole edifice that Broecker elegantly erected to explain it has collapsed like a house of cards. At the same time the climatic importance of the event seems to be reinforced, as it

MHC PROTEIN STRUCTURE

Getting into the groove

Peter Parham

In October 1968, purified papain-solubilized HLA-A2 made its debut in the pages of *Nature*¹; nineteen years later the crystallographic structure of this human class I major histocompatibility complex (MHC) molecule provided dramatic insights into T-cell biology and laid the ground rules for future experiments and speculation^{2,3}. Perhaps of greatest impact was the 'extra electron density' on display — the hazy, ill-defined image of something filling the binding site, with far-reaching implications for antigen presentation, MHC restriction and T-cell tolerance. From the genetic perspective it was their diversity and polymorphism that drew attention to MHC proteins; and although modelling of many MHC sequences into the A2 structure indicated that residues lining the sides and floor of the peptide-binding groove were the foci for advantageous amino-acid substitution^{2,4}, little was revealed of the architectural details resulting from this diversification of the antigen recognition site. With refinement of HLA-A2 and solution of the structure of HLA-Aw68, Garrett *et al.*⁵ now begin to provide such insight: on page 692 of this issue they guide us through the interstices of the celebrated 'groove', and in so doing complete what is arguably the first crystallographic experiment in MHC genetics.

Polymorphic pockets for peptides. It is the extracellular region of class I molecules that binds peptides, T-cell receptors and either the CD4 or CD8 coreceptor⁶. These molecules consist of two immunoglobulin domains and two 'peptide-binding' domains. For HLA-A2, HLA-Aw68 and other class I MHC molecules, the polymorphic and membrane-anchoring

did, indeed, almost totally halt the dramatic retreat of the great ice sheets. Doubtless Broecker will reconstruct his hypothesis even more elegantly in his next contribution. □

N. J. Shackleton is in the Godwin Laboratory for Quaternary Research, University of Cambridge, Free School Lane, Cambridge CB2 3RS, UK.

1. Fairbanks, R.G. *Nature* **342**, 637–642 (1989).
2. Mesolella, K.J., Matthews, R.K., Broecker, W.S. & Thurber, D.L. *J. Geol.* **77**, 250–274 (1968).
3. Matthews, R.K. *Quat. Res.* **3**, 147–153 (1973).
4. Bloom, A.L., Broecker, W.S., Chappell, J.M.A., Matthews, R.K. & Mesolella, K.J. *Quat. Res.* **4**, 185–205 (1974).
5. Lighty, R.G., McIntyre, I.G. & Stuckenrath, R. *Coral Reefs* **1**, 125–130 (1989).
6. Fairbanks, R.G. & Matthews, R.K. *Quat. Res.* **16**, 181–196 (1978).
7. Shackleton, N.J. *Quat. Sci. Rev.* **6**, 183–190 (1987).
8. Broecker, W.S. *et al.* *Nature* **341**, 318–320 (1987).
9. Broecker, W.S., Peteet, D.M. & Rind, D. *Nature* **315**, 21–26 (1985).

heavy chain contains both peptide-binding domains (α_1 and α_2) and one immunoglobulin domain (α_3), while a second, and in humans invariant, polypeptide (β_2 -microglobulin) provides the second immunoglobulin domain⁷. In contrast,

class II molecules have a symmetrical, and probably more ancient, arrangement in which two membrane-spanning polypeptides each contributes one peptide-binding and one immunoglobulin domain.

With minor perturbations, the polypeptide backbones of HLA-A2 and HLA-Aw68 are extremely similar⁵. This greatly simplifies comparison of the two structures for the differences between them are primarily due to the nature of amino-acid side chains at 13 positions (six in α_1 , six in α_2 and one in α_3) of substitution in the heavy-chain sequences (Fig. 1a). Ten of these residues are at positions lining the floor and sides of Garrett's groove and the differences in their chemistry, size and polarity have profound effects on its contours and on the peptides that bind, as is vividly shown by differences in the extra electron density of the two crystals.

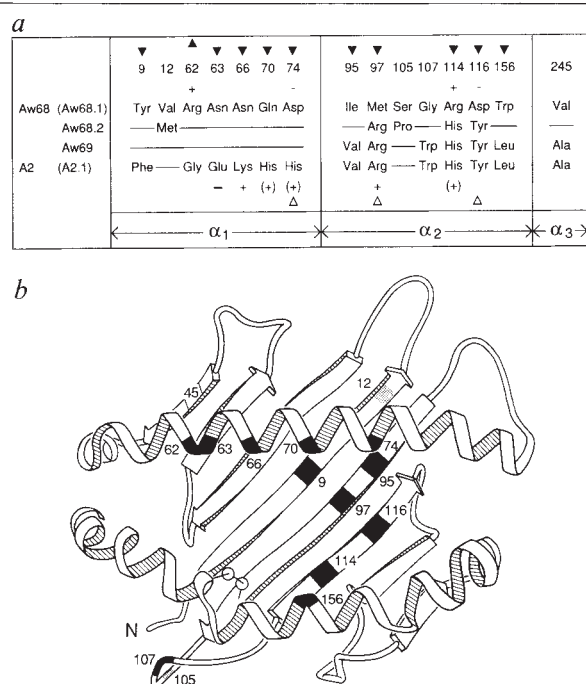
Although the antigen-recognition site (α_1 plus α_2) is symmetrical, the disposition of substitutions between A2 and Aw68 is not. A 'boomerang' pattern of variation, extending along the α_1 helix and then traversing the β structure of α_2 to the centre of the α_2 helix (Fig. 1b), is in fact similar to the distribution of high-variability residues obtained from accumulated HLA-A,B,C sequences⁷. The reason for such segregation of variation and conservation within the binding groove is not discerned by Garrett *et al.*⁵

FIG. 1 a, Amino-acid substitutions between A2, Aw68 and related molecules. Sequences are compared to Aw68.1 and only positions of difference are shown: identities with Aw68.1 are given by lines. The ten substitutions between A2.1 and Aw68.1 that line the peptide-binding groove are indicated by (▼) and the substitution at 62 that is hypothesized to affect T-cell receptor contacts is given by (▲). Residues affecting the 74 pocket are indicated by (Δ). Charged residues are noted by plus or minus. Histidines, a characteristic of A2, are indicated with (+) as their protonation under physiological conditions is uncertain.

The nomenclature for human class I MHC molecules is confusing because names have traditionally been given on the basis of serological reactivities which often encompass the products of more than one allele. Thus serologically defined HLA-A2 is known to include at least nine distinct proteins of which the 'A2' structure^{2,3} derives from HLA-A2.1.

Serologically defined HLA-A28 was initially divided into HLA-Aw68 and HLA-Aw69, and subsequently HLA-Aw68 into HLA-Aw68.1 and Aw68.2 (Garrett *et al.*'s structure is of Aw68.1). Revision of nomenclature in terms of sequence is fortunately in progress.

b, Schematic representation of the α_1 and α_2 domains of the HLA-A structure showing the 'boomerang' pattern of substitutions. Differences between A2.1 and Aw68.1 are shaded solid, substitutions specific to Aw68.2 are shaded with stipple. Position 45, which is invariant in HLA-A sequences and contributes to the conserved 45 pocket, is also shown. (Adapted from Fig. 2 of Bjorkman *et al.*².)



and must await the crystallization of homogeneous complexes in which 'extra electron density' will be transformed into high-resolution peptide antigen.

Garrett *et al.* map out features — pockets, clefts, ridges, intrusions and depressions — which impart both shared and distinctive characters to the antigen-binding grooves of A2 and Aw68. In particular, the authors draw our attention to two pockets extending under the helix of the α_1 domain, in which side chains or the ends of peptides can reside. The 45 pocket — residues 24, 26, 34, 45 and 67 — is shared by both proteins and from inspection of sequences would seem to have been conserved in all HLA-A molecules except A1. In contrast, this pocket has undergone considerable diversification in HLA-B molecules, which exhibit at least 13 structural motifs⁷. Unlike the 45 pocket, the 74 pocket of Aw68 is absent in A2 due to critical substitutions at positions 74, 97 and 116, and, as is graphically shown in Fig. 4b of Garrett *et al.*'s paper (page 693), it is packed with 'extra electron density'.

In addition to point mutation, various recombinational events, which exchange segments of homologous sequence between genes, contribute to class I MHC heavy-chain evolution. We now see that the three-dimensional consequence of such exchange is selectively to introduce, delete or modify particular pockets or sub-sites of the binding groove. For example, sequence exchanges involving A2 and Aw68 genes have produced two additional HLA-A molecules, Aw68.2 and Aw69, with antigen recognition sites that are composites of those found in A2 and Aw68 (ref. 8, Fig. 1). Aw69 is the result of a simple recombination and has the α_1 domain of Aw68 combined with the α_2 domain of A2. In Aw68.2, two pairs of inserted residues and one isolated substitution are found in a structure that is otherwise identical to Aw68. The effects of these substitutions on the antigen-binding site are localized and particularly affect the 74 pocket. For although Aw68.2 and Aw69 retain the aspartic acid of Aw68 at position 74, they both have A2-like residues at 97 (arginine) and 116 (tyrosine) that restrict access to the pocket and change its electrostatic properties (Fig. 1).

Aw68: functional or fossilized? In their discussion, Garrett *et al.* concentrate on ten differences in the peptide-binding groove and their effects on interactions with antigen, as visualized by the extra electron density⁵. There are, however, an additional three substitutions that contribute to interactions with proteins not present in the crystals; 107 with alloantibodies, 62 with T-cell receptors and 245 with the CD8 glycoprotein. Of particular interest is the substitution at position 245.

Class I MHC molecules are bifunc-

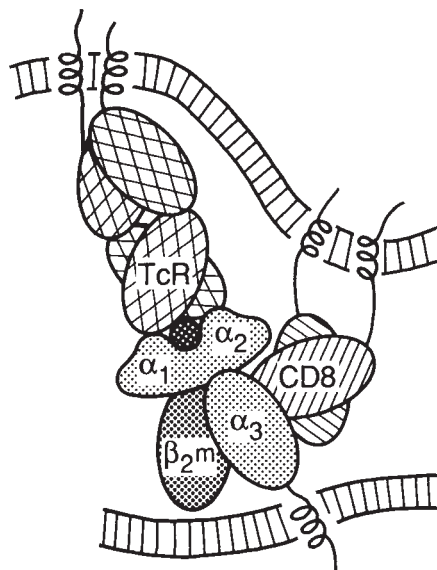


FIG. 2 Cartoon showing simultaneous interactions of both T-cell receptor and CD8 co-receptor with the complex of peptide bound to a class I MHC molecule. It is hypothesized that such complexes are important for efficient positive and negative selection of T cells in the thymus⁶.

tional: peptide-binding and T-cell receptor interactions are provided by the α_1 and α_2 domains, while CD8 interacts with α_3 (refs 9, 10). There is increasing evidence that CD8 has a critical role in positive and negative selection of class I restricted T cells in the thymus⁶, perhaps involving the simultaneous binding of both CD8 and T-cell receptor to the same complex of self-peptide and class I MHC molecule (Fig. 2). If true, this raises the possibility that Aw68, unlike A2, is functionally deficient. For, despite their structural similarity, Aw68 has considerably reduced affinity for CD8 compared to A2 and other class I MHC molecules, a property resulting from the substitution of valine for alanine at position 245 in the α_3 domain⁹. It has also been difficult to obtain T cells specific for Aw68, although some T cells stimulated by A2 or Aw69 crossreact on Aw68.1, providing circumstantial evidence that Aw68 may be inefficient in positive thymic selection⁹.

There is wide variation in the number of class I genes of different species and also in their structure, diversity and expression. These observations, together with the identification of pseudogenes, which appear to be remains of once-functional antigen-presenting genes, suggest an evolutionary dynamism for the class I genes, with continuing formation, diversification and degeneration of alleles and loci^{11,12}. This is probably due to the changing demands upon antigen presentation and T-cell immunity imposed by a variable antigenic environment. At any point in time, MHC alleles at various stages of ascendancy and decline will exist, and Aw68 may represent a molecule that has been recently retired by mutation in the

CD8 binding site. One advantage of such a lay-off could be the release from Aw68-mediated thymic deletion of receptors that can be positively selected by the other MHC molecules of an individual (A2, for example). Thus one can think of Aw68, not as inactive, but as an MHC molecule emeritus — one quite capable of binding and presenting antigen but no longer able to impose its will on the T-cell repertoire.

More MHC structures. Since 1987 immunologists have exuberantly exploited HLA-A2 as a structural model for other MHC molecules. Although the similarity of A2 and Aw68 will, for some, renew faith in these exercises of extrapolation, it is clear from the comparisons of Garrett *et al.* that small differences in structural details are very important — specially for T cells. Structural generalizations might also be tempered by the knowledge that A2 and Aw68 were chosen for investigation precisely because they share physico-chemical properties that distinguish them from other HLA-A,B molecules and which facilitated their purification¹. From nucleotide sequences it is now appreciated that HLA-A alleles form five families with A2 and Aw68 (also Aw68.2 and Aw69) being members of the same family^{8,13}.

Thus, even within HLA-A molecules, the known structures cover a limited territory, and one wonders to what extent the more diverse and more rapidly diverging HLA-B molecules will conform to the A2/Aw68 plan. In the case of class II MHC molecules, the biological differences — in antigen presentation, intracellular traffic, invariant chain association and in binding to CD4 rather than CD8 — surely demand reappraisal of the class I blueprint. At the recent Cold Spring Harbor Symposium, Don Wiley described crystals for the extracellular domains of one HLA-B molecule and one class II MHC molecule, HLA-B27 and HLA-DR1, structures that promise to provide answers to these questions. □

Peter Parham is in the Department of Cell Biology, Stanford University, Stanford, California 94305, USA.

1. Sanderson, A.R. *Nature* **220**, 192–195 (1968).
2. Bjorkman, P.J. *et al.* *Nature* **329**, 506–512 (1987).
3. Bjorkman, P.J. *et al.* *Nature* **329**, 512–518 (1987).
4. Hughes, A.L. & Nei, M. *Nature* **335**, 167–170 (1988).
5. Garrett, T.P.J., Saper, M.A., Bjorkman, P.J., Strominger, J.L. & Wiley, D.C. *Nature* **342**, 692–696 (1989).
6. Janeway, C.A. Jr *Nature* **335**, 208–210 (1988).
7. Bjorkman, P.J. & Parham, P. *A. Rev. Biochem.* (in the press).
8. Holmes, N., Ennis, P., Wan, A.M., Denney, D.W. & Parham, P. *J. Immun.* **139**, 936–941 (1987).
9. Salter, R.D. *et al.* *Nature* **338**, 345–347 (1989).
10. Connolly, J.M., Potter, T.A., Wormstall, E.-M. & Hansen T.H. *J. exp. Med.* **168**, 325–341 (1988).
11. Howard, J.C. in *Evolution and Vertebrate Immunity* (eds Kelsoe, G. & Schulze, D.H.) 397–411 (University of Texas Press, Austin, 1987).
12. Klein, J. & Figueroa, F. *CRC Crit. Rev. Immun.* **6**, 295–386 (1986).
13. Kato, K., Trapani, J.A., Allopenna, J., Dupont, B. & Yang, S.Y. *J. Immun.* **143**, 3371–3378 (1989).

The Big Bang's magic number

Frank Close

INITIAL data from the LEP electron-positron collider at CERN, Geneva, now formally published in *Physics Letters*¹⁻⁴ confirm one of the cornerstones of modern physics, namely the Big Bang model for creation of matter.

The Big Bang model rests, in large part, on the ability of theory to explain the relative abundances of the light elements, in particular the primaevial ratio of hydrogen to helium. The fusing of these light elements took place during a brief interval in the early Universe. Before this time the heat was too great for the nuclei to stay bound; beyond this time free neutrons had decayed and were no longer available for the alchemy. The elemental abundances depend, *inter alia*, on the amount of time the Universe spent in these judicious conditions, which depends on its rate of expansion and this, in turn, depends on the number of light particle varieties around — hence on the number of light neutrinos. Cosmologists had thereby determined that they could explain the elemental abundances if there were at most four varieties, or 'flavours', of light neutrino, where light means less than about 1 MeV (that is less than about twice the mass of the electron).

The LEP results reveal the number of light neutrinos directly, without any assumptions about elemental abundances or cosmology. The essential key is that counter-rotating beams of electrons (e^-) and positrons (e^+) annihilate at an energy tuned such that they produce the massive Z^0 — the carrier of the neutral weak interaction. Among the various decay pathways for the Z^0 are pairs of neutrinos (precisely, a neutrino and an antineutrino). Thus, just as water escapes from a leaky bucket faster the more holes there are, so the lifetime of the Z^0 is shorter the more neutrino flavours there are for it to decay into. It is the precise measure of this lifetime that has determined the number (N_ν) of light neutrino flavours.

The quality of the LEP results may be illustrated by comparison with previous knowledge on N_ν . C. Hearty *et al.*⁵ analysed lower-energy data on $e^-e^+ \rightarrow \gamma + \text{missing neutral energy}$. If the missing neutral energy is carried by neutrinos, this gives $N_\nu < 5.2$. In that same paper, the analysis of W and Z^0 production and decay data at CERN's older proton-antiproton collider gives $N_\nu < 7.0$ for the UA1 and UA2 collaborations individually (though this involves further information from other processes and is thus somewhat indirect). The analogous Z^0 production at the Fermilab Tevatron hadron collider was studied by the CDF collaboration⁶; this gives a rather good

measure of the Z^0 mass but does not significantly limit N_ν .

The significant advances have come from the custom-designed electron-positron machines SLC (Stanford Linear Collider) and LEP. The SLC uses new concepts in collider technology and is currently accumulating data slowly. Their results are limited by the statistics; their original publication (cited in David Miller's News and Views⁷) gave $N_\nu = 3.8 \pm 1.4$; by the time of the Stanford Lepton-Photon Symposium in August this had improved to 3.0 ± 0.9 and the latest refinement⁸ is to 2.8 ± 0.6 .

LEP is a conventional circular collider that has performed to design specification in a miracle of technological success. The four large collaborations have made their first reports on the Z^0 parameters and obtain for $N_\nu = 3.42 \pm 0.48$ (L3¹), 3.27 ± 0.30 (ALEPH²), 3.12 ± 0.42 (OPAL³) and 2.4 ± 0.65 (DELPHI⁴). Thus the odds against even four flavours are drastically lengthened and the three known may be the sum total. Moreover, cosmologists defined 'light' as a mass below 1 MeV; the SLC and LEP data extend this by a factor of 40,000 to below 40 GeV.

The Z^0 decays to hadrons and charged leptons as well as to neutrinos: so the neutrino decays are only a part of the whole, which limits the accuracy of deduction of N_ν . The best bet for improving the limits on N_ν seems to be LEP, where the machine works so well that

PHASE TRANSITIONS

New ideas for the melting pot

Robert W. Cahn

MELTING is distinct from other phase transitions in that it is not clear which of several mechanisms is responsible for initiating the process as temperature is raised. On page 658 of this issue¹, Tallon joins a debate on what effect each of the possible precursors to melting may have on the extent to which crystals can be superheated.

Following my discussion² three years ago of the first observations of crystalline superheating above the equilibrium melting point, and the intimate linkage of this phenomenon with the nature of the crystal surface, Maddox³ described new, strong experimental evidence that melting is indeed initiated at a crystal surface. Last year, Fecht and Johnson⁴ (with myself acting as chorus⁵) proposed that, under circumstances that permitted superheating at all, there must be an upper limit, thermodynamically determined, to

good statistics may be anticipated for $e^+e^- \rightarrow \gamma + Z^0$, where $Z^0 \rightarrow \text{neutrinos}$. Thus one will have $e^+e^- \rightarrow \gamma + \text{missing neutral energy}$ which is directly proportional to N_ν (the same process as discussed by Hearty *et al.* but now at energies where a real Z^0 leads to the neutrinos) without contamination of the other decay channels. P. Langacker⁹ has taken some of these recent data and studied their implications for the yet undiscovered Higgs boson and top quark masses. They do not constrain the former but do constrain $m_t < 210$ GeV (the top quark mass enters into radiative corrections that affect some weak-interaction processes). The weak-interaction angle (Weinberg angle θ_w) is known within errors as a function of m_z and m_t . The precise measure of m_z thus gives a rather clear linear relation between $\sin^2\theta_w$ and m_t , and it is the errors on $\sin^2\theta_w$ that allow the range of m_t up to 210 GeV.

A reduction of these errors, by a more precise measure of m_w/m_z , will thus lead to a prediction for the mass of the top quark, the missing member of the three families of the standard model. And if the SLC and LEP data are to be believed, three is the magic number. □

Frank Close is at Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA.

1. Adeva, B. *et al.* *Phys. Lett.* **B231**, 509–518 (1989).
2. Decamp, D. *et al.* *Phys. Lett.* **B231**, 519–529 (1989).
3. Akray, M. Z. *et al.* *Phys. Lett.* **B231**, 530–538 (1989).
4. Aarnio, P. *et al.* *Phys. Lett.* **B231**, 539–547 (1989).
5. Hearty, C. *et al.* *Phys. Rev.* **D39**, 3207 (1989).
6. Abe, F. *et al.* *Phys. Rev. Lett.* **63**, 720–723 (1989).
7. Miller, D. J. *Nature* **340**, 677–678 (1989).
8. Abrams, G. S. *et al.* *Phys. Rev. Lett.* **63**, 2173–2176 (1989).
9. Langacker, P. *Phys. Rev. Lett.* **63**, 1920–1921 (1989).

how far superheating can go. The central idea was that an 'entropy catastrophe', beyond which the crystalline entropy would exceed that of the liquid phase, would determine this limit. A few months later, Lele *et al.*⁶ proposed an alternative way of estimating the temperature of the upper entropy catastrophe (there is also a lower catastrophe, the Kauzmann temperature, that marks the limit of supercooling of a liquid/glass). It is this same question — the best way to estimate the upper limiting temperature for superheating — that Tallon also addresses.

The crucial role of the surface in initiating melting has been underlined in several recent papers. The important work by Pluis *et al.* to which both Maddox³ and I² drew attention has now been more fully described⁷: as the melting point is approached, a layer of disordering gradually develops; its thickness grows very

rapidly within a few degrees of the melting point. Pontikis and Sindzingre⁸ reviewed the tenuous evidence linking this surface roughening of crystals to surface melting and concluded that the two phenomena are probably entirely distinct. A bevy of theorists headed by Lipowsky (who has done earlier distinguished work in this field) has pursued the analysis of surface melting, and its anisotropy (which is a feature of the results from Pluis *et al.*) in terms of critical theory⁹. They introduce a family of novel order parameters to govern the melting process which depend on features of both crystal and liquid — thus beginning to address the old objection that most theories of melting look at the free energy of the crystalline phase but ignore that of the liquid altogether. Lipowsky's approach to melting readily interprets the role of the surface in nucleating the process.

Stoltzke *et al.*¹⁰, who specialize in computer molecular dynamics simulations, added to the growing evidence that the surface layer acquires a very high concentration of vacancies (several per cent) just before it melts — a feature that Fecht and Johnson⁴ took into account in their theory of an upper entropy catastrophe.

The main innovation in Tallon's new treatment is that he reintroduces into the analysis of upper and lower limiting temperatures for the crystalline and liquid/gas states, respectively, one of the oldest ideas in discussions of melting — Born's postulate¹¹ of a 'rigidity catastrophe', a temperature at which a shear modulus in the crystal falls to zero so that the crystal cannot sustain itself. This notion has been discounted of late, because real crystals do not lose their rigidity completely at the melting point. Using recent Monte Carlo calculations as a starting-point, Tallon shows that Born's idea is correct after all if the temperature of zero rigidity is taken to be the freezing point rather than the melting point (he argues that these points do not coincide in temperature–pressure–volume phase space). He demonstrates that an upper (absolute) instability limit for the crystalline state of aluminium, defined by the vanishing of the crystal rigidity, is at a temperature lower than the instability limits defined by an entropy

catastrophe^{4,6} and by an isochoric catastrophe (which he also treats) at which the specific volumes of crystal and liquid become equal. The rigidity catastrophe, Tallon claims, is what imposes the real upper limit to which a crystal can in principle be superheated.

There is just one feature of Tallon's treatment that is faintly disturbing. Instead of using the straightforward curve of free energy against temperature for the liquid, he subtracts the communal entropy of the liquid; this means that each atom is now constrained to remain within an unchanging set of neighbours or, in other words, the liquid is treated as a glass of infinite viscosity. It is not really clear why this is thought apposite for a temperature

NEUROSCIENCE

A finger on brain receptors

Charles F. Stevens

FOR the past several years, laboratories around the world have been racing to clone one or another of the various types of glutamate receptor present in the brain: these are a class of protein responsible for excitatory transmission between neurons. On pages 643, 684 and 689 of this issue, three groups^{1–3} now report sequences of receptors that specifically bind kainate, a well-known glutamate analogue. Hollmann *et al.*¹ demonstrate that they have, indeed, cloned a functional glutamate receptor channel. Their observations are especially striking because the new kainate receptor seems to define a new family of brain receptors. The other two laboratories have cloned something distinctly different but related, and discovering what these other proteins do may well turn out to be one of the most interesting results of the race.

The proteins sought belong to the functionally defined class of ligand-gated channels that serve as the receptors for neurotransmitters at synapses. Members of three distinct families of ligand-gated channels have been cloned so far, and these all belong to a superfamily. As well as homology of sequences, a common structural characteristic exhibited by all of the family members is a pattern of four membrane-spanning helices. This pattern is detected by a hydrophobicity plot that reveals four regions of the protein long enough and hydrophobic enough to form an α -helix that stretches through the core of the lipid bilayer. Because the ligand-gated channels possess a specialized receptor site for their ligand, they have pharmacological properties defined by various chemicals that either block the binding site or act like the natural ligand. The operation of ligand-gated channels can be detected electrophysiologically because application of the appropriate ligand causes the channel to open and

range in which the liquid in fact is bound to be highly mobile.

We now have three papers^{1,4,6} that deal with the behaviour of crystals superheated through hundreds of degrees until they come up against a 'no further' sign. Perhaps the time is now ripe for experimentalists, following the pioneers^{12,13}, to explore further the conditions under which various crystalline species can be metastably superheated, so that an experimental test of the various theories of the upper absolute instability limit will perhaps become feasible. □

Robert W. Cahn is in the Department of Materials Science and Metallurgy, University of Cambridge, Cambridge CB2 3QZ, UK.

permits the transmembrane flux of an ionic current.

What all of the laboratories expected to find, then, was a protein that opens a pore when glutamate is bound. This receptor should have the pharmacological properties already defined by studies on brain glutamate receptors; in the cases at hand, for example, the glutamate analogue kainate must bind with high affinity and produce channel opening. Further, the receptor was expected to have sequence homologies that identified it as a member of the ligand-gated-channel superfamily and to have the characteristic four-membrane-spanning hydrophobicity plot.

What, then, has been found? Hollmann *et al.*¹ have cloned what must be at least part of a glutamate receptor channel, because injection of message made from their clones into frog oocytes endows the oocyte membrane with electrophysiological properties typical of neurons that contain kainate receptor channels. This observation was an inevitable result of the cloning strategy adopted: the screen Hollmann *et al.* used to identify the appropriate clone was expression of function (electrophysiological response to applied kainate) in oocytes. The protein found by this procedure has a relative molecular mass of about 100,000 and a hydrophobicity plot similar to that of members of the ligand-gated-channel gene family. Unexpectedly, however, the homology between the kainate receptor and members of the channel superfamily is at most extremely weak — so weak, in fact, that for practical purposes these clones define a new family of ligand-gated channels.

The other two groups adopted the more usual alternative to the expression-cloning approach used by Hollmann *et al.*: they identified a protein that binds kainate with high affinity, purified it, determined the amino-acid sequence of fragments, made

1. Tallon, J.L. *Nature* **342**, 658–660 (1989).
2. Cahn, R.W. *Nature* **323**, 668–669 (1986).
3. Maddox, J. *Nature* **330**, 599 (1987).
4. Fecht, H.J. & Johnson, W.L. *Nature* **334**, 50–51 (1988).
5. Cahn, R.W. *Nature* **334**, 17–18 (1988).
6. Lele, S., Ramachandrarao, P. & Dubey, K.S. *Nature* **336**, 567–568 (1988).
7. Pluis, B., Frenken, J.W.M. & van der Veen, J.F. *Phys. Scr.* **T198**, 382–386 (1987).
8. Pontikis, V. & Sindzingre, P. *Phys. Scr.* **T198**, 375–381 (1987).
9. Lipowsky, R., Breuer, U., Prince, K.C. & Bonzel, H.P. *Phys. Rev. Lett.* **62**, 913–916 (1989).
10. Stoltzke, P., Nørskov, J.K. & Landman, U. *Surf. Sci.* **220**, L693–L700 (1989).
11. Born, M. *J. chem. Phys.* **7**, 591–601 (1939).
12. Däges, J., Gleiter, H. & Perepezko, J.H. *Mat. Res. Soc. Symp. Proc.* **57**, 67–78 (1987).
13. Rossouw, C.J. & Donnelly, S.E. *Phys. Rev. Lett.* **55**, 2960 (1985).

corresponding oligonucleotide probes, and screened brain cDNA libraries by hybridization. Wada *et al.*² and Gregor *et al.*³ have cloned what may well be the frog and chick versions of the same kainate-binding protein (KBP). This protein has a relative molecular mass of about 50,000 (about half that of the kainate receptor channel identified by Hollmann *et al.*) and, again, the hydrophobicity plot typical of the ligand-gated-channel gene family. The KBP exhibits about 25 per cent amino acid similarity to the last (carboxy) half of the kainate receptor channel cloned by Hollmann *et al.*, but the first 350 amino acids of the kainate receptor channel are entirely absent from the KBP. The KBP binds kainate, as it should, but does not function as a channel when expressed in oocytes.

In summary, then, Hollmann *et al.* have cloned at least one subunit of a functional glutamate receptor channel; the other two laboratories have cloned a related protein that binds glutamate (kainate) but whose function is at present unknown. What could this KBP be? One possibility is that it is a subunit of a neuronal kainate receptor and that the channel will only function when other subunits are present. Because the chick KBP is remarkably abundant in cerebellum and is only present on Bergmann glial cells there, other possibilities are likely. The KBP could, for example, be a component of an uptake system or part of a glutamate receptor involved in a signalling that does not use an ion channel. Glia have recently been shown to transmit signals to one another in response to glutamate⁴, and this signalling system, which involves changes in intracellular calcium concentration, responds selectively to kainate. Whatever the KBP

is, it belongs to a previously undefined family; and because of its abundance in glia rather than neurons, it must play a role in some as yet unknown aspect of brain function.

For neurobiologists, the cloning of a glutamate receptor channel is an exciting event. To understand why, one must appreciate the role of glutamate receptor channels in brain function. First, glutamate is the main excitatory neurotransmitter in the brain, and every neuron — as far as we know — has ion channels with receptors for glutamate. The glutamate receptors are thought to form a complex family because they fall into a rich variety of pharmacological types, and, if past experience is any guide, the gene families will be even richer than the current pharmacology indicates. Second, at least one pharmacologically defined type of glutamate receptor, the one named after the glutamate analogue NMDA (*N*-methyl-D-aspartate) that best activates it, is especially interesting. The NMDA receptor is a site of action of the street drug, 'angel dust', is responsible for excitotoxic nerve cell death and has a key part in long-term potentiation, the phenomenon thought to underlie some forms of memory. Furthermore NMDA receptors have been implicated in certain types of seizure and

various neurodegenerative diseases, and in the reorganization of neuronal connections that depends on experience. Heinemann's laboratory, from which the paper by Hollmann *et al.* comes, has already started finding other family members related to their kainate receptor channel, and it is expected that an entirely new family of receptors, one that may include the NMDA receptor, will be identified through homology screening.

Finally, an interesting sociological note. Although the three groups that came first in the race to clone a glutamate receptor worked independently (and competitively), the research of two of them was carried out in different laboratories at the Salk Institute, and one of the co-authors in the third group was in yet another laboratory at the Salk. Something about southern California must be good for research on excitatory amino acids. □

Charles F. Stevens is in the Section of Molecular Neurobiology, Yale Medical School, 333 Cedar Street, New Haven, Connecticut 06510, USA.

1. Hollmann, M., O'Shea-Greenfield, A. Rogers, S. W. & Heinemann, S. *Nature* **342**, 643–648 (1989).
2. Wada, K. *et al.* *Nature* **342**, 684–689 (1989).
3. Gregor, P., Mano, I., Maoz, I., McKeown, M. & Teichberg, V. I. *Nature* **342**, 689–692 (1989).
4. Cornell-Bell, A., Finkbeiner, S. M., Cooper, M. S. & Smith, S. J. *Science* (in the press).

BIOGEOCHEMISTRY

Effects of increased ultraviolet

T.E. Graedel

THE discovery of the Antarctic ozone hole has sparked tremendous discussion about the implications of increased ultraviolet radiation on geochemical and biochemical processes. Because many of the atmospheric and marine chemical cycles strongly interact with those of the biota, changes in any of the systems imply changes in all. These interlocking concerns formed the basis for a recent workshop* which brought together atmospheric chemists and physicists, marine and freshwater chemists, biochemists and biologists.

The effects of changing ultraviolet fluxes can be studied by examining each term of what one might call the photochemical impact equation. For a specific molecule *i*, the impact *I_i* can be written as

$$I_i = \int_{\lambda_{\min}}^{\lambda_{\max}} J(\lambda) \sigma_i(\lambda) \phi_i(\lambda) d\lambda$$

in which *J*(λ) is the incident flux, $\sigma_i(\lambda)$ is the absorption cross section and $\phi_i(\lambda)$ is the quantum efficiency (that is, the fraction of absorbed photons that cause the process under study, as opposed to having their energy dissipated in some other

way). The integration limits are the atmospheric cut-off at about 290 nm at the short wavelength end, and the longest wavelength capable of producing an effect (generally in the range 350–400 nm) at the upper end.

In the case of absorption by agglomerated particles — poorly characterized dissolved species or organisms, for which molecule-by-molecule approaches are inappropriate — the absorption cross-section and quantum yield terms are replaced by *A_k*(λ). This 'action spectrum' is an experimentally determined measure of the effects of the interaction of light at a particular wavelength with a specific constituent. A quantitative assessment of the effects of changing ozone concentrations requires a relatively complete understanding of each term in these equations; speakers at the workshop presented results related to each term.

Changes in atmospheric ozone are now being monitored widely. It does not follow, however, that even *J*(λ), the present flux of solar radiation to the surface, is known accurately, because clouds play such a large role in the radiation balance. The distribution and physical properties of clouds are not well characterized, particularly over the oceans and

100 years ago

EARLY EGYPTIAN CIVILIZATION.

THOUGH debarred from the richest districts of Egypt — owing to national jealousies — I have been fortunate enough to discover two small towns which have revealed the works of the Middle and New Kingdoms with chronological exactness. Beside the Egyptian interest of these places, they are of prime importance for Mediterranean history, having been colonies of foreign workmen. The most novel discovery of



Signs incised on pottery of the twelfth dynasty

all is the presence of apparently alphabetic signs in use in both towns (figure) at about 2500 B.C. and 1300 B.C. The apparent connection of these signs with some of the mason's marks suggests that they may have been adopted by the foreign workmen from their Egyptian fellow-labourers; and the very lack of literary education among such men would lead to their forming alphabets of their own.

From *Nature* **XLI**, 109–111 (1889).

*Effects of Solar Ultraviolet Radiation on Biogeochemical Dynamics in Aquatic Environments, Woods Hole, Massachusetts 23–25 October, 1989.

the poles. Calculations by S. Madronich (National Center for Atmospheric Research, Boulder) predict that for mid- and low-latitudes, fluxes of biologically active UVB (290–320 nm) at the surface will change by –5 to +10 per cent by the year 2060, depending both on the model and the assumed course of ozone depletion. For southern polar regions, large increases in springtime ultraviolet are expected (as much as a factor of ten in the year 2060). But because the average UVB irradiance is much higher at equatorial latitudes than at the poles, a percentage increase in UVB irradiance in the tropics corresponds to many more additional photons than the same increase at high latitudes.

Increases in solar ultraviolet radiation seem certain to produce substantial changes in the composition of the atmosphere as a consequence of the absorption of the radiation by gaseous molecules. Calculations by A.M. Thompson (NASA/Goddard Space Flight Center) indicate that the oxidizing power of the lower atmosphere will be much increased and that the atmospheric content of many trace gases scavenged by oxidizers will decrease. Cloud effects pose large uncertainties, however, because soluble gas-phase species can interact with clouds (and aid in their formation) in ways not well understood, and because most atmospheric chemical models have yet to treat cloud effects.

P. Crutzen (Max-Planck-Institute for Chemistry) reported recent calculations in which cloud chemistry was added to a global two-dimensional model. Although clouds occupy only about 15 per cent of the volume of the lower troposphere and liquid water occupies only 10^{-7} to 10^{-6} of a cloud's volume, the inclusion of clouds in the model greatly decreased the calculated concentrations of ozone, hydroxyl radical and nitrogen oxides, thereby affecting the budgets of many other trace gases (P. Crutzen and J. Lelieveld *Nature*, in the press).

Within condensed phases, the results of increased ultraviolet are also likely to be dramatic. A few years ago, there was little evidence for the existence of reactive chemical transients in environmental waters. Recently, however, several studies using carefully designed trapping-agent molecules have established the presence of such species as excited states, singlet molecular oxygen, solvated electrons, hydroxyl radicals, superoxide radicals, carbon-centred radicals and hydrogen peroxide. All are generated by solar ultraviolet radiation. Hydrogen peroxide is the longest lived of these transients and its widespread occurrence in the upper ocean and freshwaters provides an important marker of the influence of solar ultraviolet radiation in aquatic environments.

The biological effects, on both plants

and animals, of increased ultraviolet radiation can be assessed by species only if each $A_k(\lambda)$, the appropriate species action spectrum, has been determined. Useable action spectra are, however, rare. T. Coohill (Western Kentucky University) highlighted some of the problems: uncertainty in the spectral purity of the incident radiation; difficult dosimetry determinations; the extent of biological repair before assay; the physical state of the target molecules; and response *in vivo* versus response *in vitro*. Hence, assessments of the effects of changing ultraviolet radiation are, at present, much more likely to be qualitative than quantitative.

In one of the workshop's more provocative presentations, F.M.M. Morel (Massachusetts Institute of Technology) argued that one could logically defend any of the following positions. Increased ultraviolet radiation will: promote a dramatic increase in new primary biotic productivity; practically destroy biotic productivity; or have no discernible effect on biotic productivity. His claim can be made because the quantitative information to determine the dominant effects is generally unavailable and because different factors seem likely to dominate at different places, times and seasons.

Morel's first position was supported by several participants who pointed out that the bioproductivity of the oceans often seems to be limited by the availability of trace nutrients such as iron, copper and manganese. Photochemical reactions involving bound forms of these elements increase their availability. Hence, regions of the oceans that are now nutrient-limited may benefit from increased ultraviolet radiation.

The converse position – that the increased damage to DNA will overwhelm any benefits in nutrient supply – was supported by several reports that UVB radiation markedly reduces growth in some species of marine phytoplankton. The biological impact is related to the number and efficiency of DNA repair systems in organisms. Field studies from Palmer Station, Antarctica (D. Karentz, University of California, San Francisco) demonstrate that there are enormous interspecies differences in sensitivity to ultraviolet.

Could increased ultraviolet fluxes have only little effect on biotic productivity? T. Mill (SRI International) argued that concentrations of marine and freshwater reactive chemical transients are too low and their lifetimes too short to have major adverse effects on the biota. It is also worth considering that marine phytoplankton already cope with reactive photoproducts and trace-element depletion. □

T. E. Graedel is at AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA.

Air of longevity

ONE of the appealing hazards of air travel is the ease with which you can get drunk. The reduced pressure and oxygen content of the cabin slows the metabolic elimination of alcohol, giving two drinks almost the effect of three. Daedalus therefore began to design a low-pressure pub. Drinkers would have to enter by an air-lock, but could then get drunk more cheaply, and stay that way longer. Conversely, increased pressure should speed up alcohol metabolism. A high-pressure bedroom, for example, might enable the alcoholic to burn off his entire hangover overnight, and greet the morning after with a clear head.

Daedalus is now extending this whole principle. He points out that if fed on restricted diets, nearly all animals mature more slowly and live longer than usual. Every creature studied seems to step through its whole 'life programme', from growth and sexual maturity to senescence and death, by the ticking of its metabolic clock. Animals raised in low pressures should therefore have longer, slower biological careers. If the principle works with man as well, then by spending your whole life in a cruising aircraft you might live to be a hundred. Or perhaps it would only seem like that.

Low pressure may even space out the minor milestones of life. The irregular menstrual cycle of many airline stewardesses has been attributed to their chaotic shuttling between time zones. Daedalus suspects instead that they 'age' more slowly in the air, and get out of step with ground-level time.

So DREADCO's physiologists are setting up a variable-metabolism unit. To avoid the problems of pressurized construction, Daedalus contemplated slowing the oxygen metabolism of the inmates by introducing a little carbon monoxide into the air of the building. But wiser counsels prevailed, and instead the oxygen concentration of its air will be varied directly. Daedalus expects that his under-oxygenated volunteers will lose appetite and need only two meals a day. To study ageing itself, he is appealing for middle-aged identical twins, one to live in the low-oxygen laboratory and one outside while microscopic examination compares the rate at which grey hairs appear on their scalps.

The effects of oxygen-enriched air should also be interesting. Metabolism may speed up sharply, so that obese volunteers are swiftly 'burnt down' to normal weight. Animals may grow faster: an oxygenated poultry house, for example, might hurry chickens to earlier maturity. It would be hard to extend such atmospheric modifications to free-range life, whether in chickens or man; though an oxygen boost on the air conditioning might be welcome in many crisis-prone offices. David Jones

Pneumologists breathe easy

SIR—I have read with interest the article “Model Lung Breathes True”¹ and the quoted paper by two physicists, Bonsignori and Salvini, *Il Nuovo Cimento*², which is said to “contain an intriguing calculation of gas transport in the human lung of a kind that would normally have been found in a physiological journal”. In fact, the results of similar calculations belong to standard physiological textbooks³. In my opinion, the paper of Bonsignori and Salvini is only relevant to the respiratory physiology of the early 1970s.

Whatever the model (in physiology as in physics), comparison with experimental observations is the crucial test. What Maddox has written¹, that “the outcome of the calculations is much what physiologists would expect”, is not true. The authors simulated a simple experiment, called single breath nitrogen washout, consisting of an inspiration of oxygen followed by an exhalation. Most of the expired-N₂ concentration profile, the so-called alveolar plateau, is not flat, as would be expected if the lung is a perfect mixing chamber. (The slope of the alveolar plateau is one of the standard indices in pneumology and is explained to medical students in elementary textbooks.)

The model of Bonsignori and Salvini simulates a flat alveolar plateau and the authors say that “in experiments, the single breath washout curves have rather flat alveolar plateaus, indicating uniform mixing of the alveolar and inspired gases”². It is well known, since the experiments of Krogh and Lindhard⁴ at the beginning of the century, that this is not the case. In other words, the lung model of Bonsignori and Salvini does not breathe true.

Maddox also quotes Bonsignori and Salvini as claiming to have improved on some of the approximations to which previous calculations have been subject. These improvements consist of: 1, using a better representation of the airway cross-section; 2, using a Taylor-Arris effective diffusion; and 3, discussing more appropriate boundary conditions. In a review on the subject⁵, several references (probably not known by Bonsignori and Salvini) are given where all these points have been studied. None of these have significantly improved agreement between the model simulations and the experimental observations.

As Maddox says, the physiological data are striking, the total cross-section of airways increasing from a few square centimetres in the trachea to more than a square metre at the last generations of the bronchial tree. A major consequence of the lung geometry is that when we inspire at constant flow, the concentration profile of the inspired gases remains quasi-

stationary, the convection flow of inspired gases towards the lung periphery being balanced by the diffusion flow of the residual (or alveolar) gases in the opposite direction⁵. This is the so-called diffusion front, equivalent to the stationary diffusion front in combustion science, when the combustion takes place in a tube with the appropriate convection velocity of gases.

A major development of models of gas mixing in the lung (not quoted by Bonsignori and Salvini) was the realization⁶ in the 1970s that one of the most fragile assumptions in previous work consists of using the symmetrical Weibel model of the lung⁶, which dates back to 1963. Not surprisingly, the first detailed asymmetrical model of the periphery of the human lung was recently published by Weibel's group⁷. Unexpected results appeared when a diffusion equation was written in an asymmetrical structure of ducts as a consequence of Fick's law of diffusion and matter balance in a branching tube: the flow of any gas in such a structure generates an inhomogeneous concentration distribution, even though the structure moves isotropically and homogeneously. This occurs in zones of the model that are subtended by branch points where the Péclet number (giving the relative contribution of convective to diffusive transport) is approximately unity. In the human lung during normal breathing, this corresponds to branch points subtending zones of the lung a few millimetres apart.

Simulated single-breath washout using these new models gives much better agreement with experimental observations than previous models⁵. This finding is of interest in that the single-breath washout, particularly when performed with tracer gases of different diffusivities such as He and SF₆, gives information on the structure of the very peripheral zones of the lung which are not accessible to traditional tests in pneumology nor even to modern imaging techniques such as positron emission tomography. To my knowledge, the description of gas transport in an asymmetrical lung model has no counterpart in physics or engineering, although it may have applications in the study of gas transport in porous media, or O₂ and CO₂ transport in some insect-built galleries, where the Péclet number may approach unity. It may also apply to the study of O₂ transport between capillaries and cells.

Physiology (which, in a way, is the physics of living beings) and physics were closely associated until the last century, and several discoveries in both fields were made by the same scientist. Boyle, Torricelli, Pascal, Hooke, Lavoisier, Laplace and Lagrange have contributed to the

early development of respiratory physiology, and both Poiseuille and Fick were physicians. The bridge is much narrower today and the paper by Bonsignori and Salvini in *Il Nuovo Cimento* has the merit, as Maddox says, of establishing a new link. It also points out the necessity for the physicists who approach physiology to learn at least its elementary basis, if they want to avoid the trap of rediscovering the wheel. Bonsignori and Salvini promise a new paper where they will pay particular attention to the question of the boundary conditions in the model. If they do not in the meantime read the published literature, respiratory physiologists need not hold their breath awaiting future issues of *Il Nuovo Cimento*.

M. PAIVA

*Institute of Interdisciplinary Research,
Free University of Brussels,
Campus Erasme,
Route de Lennik, 808,
B 1070 Brussels,
Belgium*

1. Maddox, J. *Nature* **341**, 381 (1989).
2. Bonsignori, F. & Salvini, M. *Il nov. Cim.* **11D**, 1247–1263 (1989).
3. Engel, L.A. in *Handbook of Physiology* (American Physiological Society, Bethesda, Maryland, 1986).
4. Krogh, A. & Lindhard, J. *J. Physiol., Lond.* **51**, 59–90 (1917).
5. Paiva, M. & Engel, L.A. *Physiol. Rev.* **32**, 750–796 (1987).
6. Weibel, E.R. *Morphometry of the Human Lung* (Springer, Berlin, 1963).
7. Haefeli-Bleuer, B. & Weibel, E.R. *Anat. Rec.* **220**, 401–414 (1988).

PCR origins

SIR—We noted with interest the short note, entitled “DNA amplification” (*Nature* **341**, 570; 1989) which suggests that scientists who use PCR “dust off” the *Journal of Molecular Biology* from 1971, and read the paper by “H. Gobind Khorana and colleagues”. Although the title and text of your note might suggest otherwise, the “lines of thought” and reference to experiments “in progress” in this 1971 article did not present a method that could be used to perform *in vitro* DNA amplification. The scientific community had to wait for such a method until, 14 years later, Kary Mullis invented the polymerase chain reaction (PCR).

Your readers might be interested to know that the articles written by Kary Mullis, Henry Erlich and their co-workers at Cetus describing the PCR method and its application have been cited in well over 1,000 publications since 1985. None of these publications cite the 1971 article to which you draw attention. Cetus is confident that Kary Mullis will continue to receive the recognition he deserves.

JEFFREY S. PRICE

*Cetus Corporation,
1400 Fifty-Third Street,
Emeryville,
California 94608, USA*

Similarity in membrane proteins

SIR—Adams and Pollard¹ recently discovered that the nonfilamentous myosins IA and IB of *Acanthamoeba* can bind to membranes, a finding with important implications concerning the role of these proteins in cell motility. We would like to draw attention to a 50-amino-acid domain in the sequence of the myosin-IB isozyme² that is shared with a diverse family of cyto-

two proteins of another single-celled eukaryote, *Saccharomyces cerevisiae*. The fus1 protein on the cell membrane is essential for cell fusion during mating^{3,7}; cdc25 is a regulator of the membrane associated adenyl cyclase. The SH3 motif is, therefore, shared by membrane-associated proteins of both single-celled and higher eukaryotes. This conservation

Myosin-IB	(983)	ALYDFARENPDE-----LTFNEGAUUTUI-NKSNPDWNEGEL-----NGQR-GUFPASVUELI-PR	ref. 2
Fodrin- α	(974)	ALYDgqekSPrE-----vTmkGdiliTII-NsNkDHWkUvE-----Ndrq-GfUPrAaYUkkI-dp	5
PLC148	(798)	ALFDyAqaeDE-----LTFtkAiiqnv-eKqegWIRGdyh-----hkkq-luFPsnYUEmu-s	3
v-Crk	(375)	ALFDKgnddgd-----LpFkkGdilkir-dKpeeqWNaEdm-----dGKR-GmiPupYUEKcrPs	4
c-Fgr	(84)	ALYDyeArteDd-----LTFtkGekfhiI-NntegDWEarsls-----sGkt-GciPsnYUapv-ds	9
Lsk	(67)	ALhsyepshdgd-----LgFekGeqlril-eqS-geWkqagst-ttGQe-GfiPfnfUaka-ns	10
c-Yes	(97)	ALYDyeArnted-----LsFkkGerfaiI-NntegDWEarsia--tGkn-GyjiPsnYUapa-ds	11
Hck	(62)	ALYDyeAihred-----LsFqkGdqmvUI-eea-geWkarsla--tke--GyjiPsnYUapv-ns	12
c-Src	(88)	ALYDyesrtetd-----LsFkkGerliqv-NntegDWHahst-ttGQt-GyjiPsnYUapa-ds	13
c-Abl	(68)	ALYDFuAgdnt-----LsitkGekrUlgynhHgeWcEaqtK--NGQ--GwvPsnYitpv-ns	14
fus1	(443)	viqDyeprltDE-----inislGekvkiI-athtdgWlvqkc-----ntQK-GsihuSuddkryln	6,7
cdc25	(64)	AaYDFNypikkdsqilsvqgGetiyiI-NKnsqWldGividdsHGKvrguFPanfgprlrs	15
Consensus		ALYDY-----D-----GD- Φ - Φ - Φ -----W-----Gd Φ P-----Y Φ -----	
		F F E E F	

Upper-case letters denote identity with the myosin-IB sequence. The amino-acid number of the first residue of each sequence is given in parentheses. Φ signifies hydrophobic residues. The PIR database was searched using the programs Prosearch (J. Collins and A. Coulson, University of Edinburgh) on an AMT DAP, and AMPS (G. Barton and M.J.E.S.) on a VAX 8700; the Geneproc program (Riverside Scientific, Seattle) was used to search the EMBL database.

plasmic proteins, many of which are associated with the plasma membrane (see figure). This domain, called SH3 or the A-box, was first described as a region of sequence similarity between the non-receptor protein-tyrosine kinases, the 148K phospholipase C (PLC), and the v-crk oncogene of the avian sarcoma virus CT10 (refs 3 and 4). It was subsequently found in the α -chain of the non-erythroid spectrin homologue, fodrin⁵.

Our database searches have also

revealed SH3 domains in cdc25 and fus1, indicates a basic function common to all eukaryotes but this function is still a mystery. Although SH3 is found in many proteins at the cell periphery, it is not universal; it is absent from erythroid spectrin for instance.

Adams and Pollard¹ mapped the membrane-binding region of myosin-IA to a 100K N-terminal chymotryptic fragment, but the binding region of myosin-IB has not yet been localized. Intriguingly, the SH3 domain of the IB protein is close to the protein's ATP-independent actin-binding site⁸. It will be interesting to see whether the SH3 domains of myosin-IB and other proteins play a role in binding either directly to membranes or to the submembrane cytoskeleton.

ADAM R. F. RODAWAY

MICHAEL J. E. STERNBERG

DAVID L. BENTLEY

ICRF Laboratory,
Lincoln's Inn Fields,
London, WC2A 3PX, UK

Retrocitation

SIR—In my News and Views article "Reverse Transcriptases. Retrons in Bacteria" (ref. 1) I stated "no such reverse transcriptase seemed to exist in bacteria". M. Beljanski has since called my attention to his publications on reverse transcriptase in bacteria²⁻⁵. This work was confirmed in several Soviet publications⁶⁻⁹. It will be of interest to see if these activities are related to retrons.

HOWARD M. TEMIN

McArdle Laboratory,
University of Wisconsin,
Madison, Wisconsin 53706, USA

1. Temin, H.M. *Nature* **339**, 254–255 (1989).
2. Beljanski, M. *C. r. heb. Séanc. Acad. Sci., Paris D274*, 2801–2804 (1972).
3. Beljanski, M. *C. r. heb. Séanc. Acad. Sci., Paris D276*, 1625–1628 (1973).
4. Beljanski, M. & Beljanski, M. *Biochem. Genet.* **17**, 163–180 (1974).
5. Beljanski, M., Bourgarel, P. & Beljanski, M.S. *C. r. heb. Séanc. Acad. Sci., Paris D286*, 1825–1828 (1978).
6. Romashchenko, A.G. et al. *Dokl. Akad. Nauk SSSR* **233**, 734–737 (1977).
7. Lushnikova, T.P. et al. *Molec. Biol.* **12**, 1163–1171 (1978).
8. Vorob'eva, N.V. et al. *Molec. Biol.* **16**, 977–983 (1982).
9. Grabkina, O.A. et al. *Molec. Biol.* **17**, 830–838 (1983).

1. Adams, R. & Pollard, T. *Nature* **340**, 565–568 (1989).
2. Jung, G., Korn, E. & Hammer, J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 6720–6724 (1987).
3. Stahl, M., Ferez, C., Kelleher, K., Kriz, R. & Knopf, J. *Nature* **332**, 269–272 (1988).
4. Mayer, B., Hamaguchi, M., & Hanafusa, H. *Nature* **332**, 272–274 (1988).
5. Wasenius, V. et al. *J. cell. Biol.* **108**, 79–93 (1989).
6. Truehart, J., Boeke, J. & Fink, G. *Molec. cell. Biol.* **7**, 2316–2328 (1987).
7. McCaffrey, G., Clay, F., Kelsay, K. & Sprague, G. *Molec. cell. Biol.* **7**, 2680–2690 (1987).
8. Lynch, T. et al. *J. Biol. Chem.* **261**, 17156 (1986).
9. Katamine, S., et al. *Molec. cell. Biol.* **8**, 259–266 (1988).
10. Marth, J., Pect, R., Krebs, E. & Perlmutter, R. *Cell* **43**, 393–404 (1985).
11. Sukeyawa, J. et al. *Molec. cell. Biol.* **7**, 41–47 (1987).
12. Klesz, M., McKercher, S. & Maki, R. *Nucleic Acids Res.* **15**, 9600 (1987).
13. Takeya, T. & Hanafusa, H. *Cell* **32**, 881–890 (1983).
14. Shtivelman, E., Lifshitz, B., Gale, R., Roe, B. & Canaan, E. *Cell* **47**, 277–284 (1986).
15. Broek, D. et al. *Cell* **48**, 789–799 (1987).

What are males good for?

SIR—Successful rearing of young in most birds requires the care of two adults. Such parental pairs are almost always a male and a female, but the discovery of female/female pairings in one tern, one goose and several gull species, discussed by Jared Diamond¹, demonstrates that other parental pairing combinations are possible. Female/female pairs are, nevertheless, rare even in these few species²⁻⁶. Determining the reasons why they are so uncommon may help to formulate answers for the old question 'What are males good for?'

The most obvious benefit to females of pairing with a male — access to a source of sperm — is probably not of primary importance, as females paired together still copulate with unmated and mated males⁷. Female/female pairs, therefore, produce fertile eggs, although their fertility rates may be less than those of male/female pairs⁴⁻⁶.

Perhaps a more important advantage provided by males is that of a territory. Female birds are generally smaller than males, and may, therefore, be at a disadvantage when competing against a male for a territory. In all the colonies where I have found female/female pairs or the unusually large (5–7 egg) clutches that result from them^{2,8} (eight ring-billed gull colonies, four California gull colonies and a Caspian tern colony), space for more nesting pairs was available on the nesting islands. In western gulls, female/female pairs were found in one colony where there was little competition for breeding space, but no such pairs were found in another colony where breeding space was limited⁹.

Probably the most important advantage a female derives from pairing with a male is the opportunity to raise only her own young (nest parasitism aside). By contrast, each female in a homosexual pair will contribute, on average, only half the young. An unrelated pair of females must, therefore, produce twice the offspring of a heterosexual pair before the reproductive output of each homosexually paired female equals that of a

1. Diamond, J.M. *Nature* **340**, 101 (1989).
2. Conover, M.R. *Condor* **86**, 467–471 (1984).
3. Conover, M.R. *Wilson Bull.* **96**, 249–267 (1984).
4. Hunt, G.L. Jr & Hunt, M.W. *Science* **196**, 1466–1467 (1977).
5. Kovacs, K.M. & Ryder, J.P. *Auk* **100**, 658–669 (1983).
6. Quinn, T.W. et al. *Auk* **106**, 177–184 (1989).
7. Fitch, M.A. & Shugart, G.W. *Am. Nat.* **124**, 116–126 (1984).
8. Conover, M.R. *Condor* **85**, 346–349 (1983).
9. Pierotti, R. *Auk* **98**, 532–549 (1981).
10. Lagrenade, M. & Mousseau, P. *Auk* **100**, 658–669 (1983).
11. Hunt, G.L. Jr et al. *Auk* **97**, 473–479 (1980).
12. Conover, M.R. & Hunt, G.L. Jr *Condor* **86**, 472–476 (1984).
13. Conover, M.R. & Hunt, G.L. Jr *Wilson Bull.* **96**, 619–625 (1984).

heterosexual female. Sister/sister pairs must produce 33 per cent more young per pair for the same result. Although female/female pairs produce large clutches of eggs there is, however, no evidence that female/female pairs produce more young than male/female pairs, and they may in fact produce substantially less^{4,5,10}. Hence, males seem to provide such overwhelming advantages to their female mates that female/female pairings occur primarily when females are unable to pair with males¹¹⁻¹³.

MICHAEL R. CONOVER

Connecticut Agricultural
Experiment Station, New Haven,
Connecticut 06504, USA

Embryo and seed abortion in plants

SIR—In asking “why do plants produce so many more ovules than seeds”, Charlesworth¹ and Wiens *et al.*² have addressed an important issue. Wiens *et al.* attribute the high level of ovule abortion observed in many plants to genetic load, a factor also emphasized by Charlesworth in her commentary. We wish to point out that data presented by Wiens *et al.* are not consistent with their own hypothesis, and briefly we review alternative explanations³⁻⁶ for ovule abortion.

Ovule abortion in plants occurs at two levels. In many species most flowers do not set fruits. Furthermore, in a number of species, all ovules within a flower do not set seeds. As Charlesworth points out, inbreeders generally have higher fruit/flower and seed/ovule ratios, but exceptions are known. For example in the Mimosoideae, regardless of the breeding system, fruit/flower ratios are extremely low and seed/ovule ratios are very high⁷. In other species with negatively skewed distribution of seeds within fruit⁸, the proportion of flowers and developing fruits that are aborted is very high as compared to the proportion of ovules aborted within the retained mature fruits. The exceptions provide insights into causes of ovule abortions and emphasize the need to consider separately the two levels of abortion, at least under some circumstances.

Wiens *et al.* deal with ovule abortion at the level of flowers. Only 2.5 per cent of flowers in *Dedeckera eurekaensis* set fruits, though many more initiate the development of ovules. Such a low level of fruit set is not unusual in plants^{8,9}. Wiens *et al.* attribute the low reproductive output in *D. eurekaensis* to segregational load but also state that “Intraplant self-pollination (geitonogamy) may be common in *D. edeckera*, but inbreeding depression resulting from selfing in a normally outcrossed population is not a major factor

contributing to the low seed set.” As Charlesworth points out the segregational load hypothesis would predict strong inbreeding depression following selfing in outcrossing populations. Wiens *et al.* did not find the difference in fruit set between self- and cross-pollinated flowers to be statistically significant. The segregational load hypothesis also predicts abortions to occur randomly along the inflorescences. It is not known if the flowers are aborted at random in *D. eurekaensis*, but in other species flower abortions are nonrandom (ref. 3 and J. H. Beach, personal communication). It is therefore premature to assume that segregational load is a major factor underlying abortion in *D. eurekaensis* (and other species).

The abortion of ovules within fruits is briefly considered by Charlesworth¹. Many species abort a variable number of their ovules after fertilization within developing fruits. Again the hypothesis of segregational load would predict random patterns of abortion. Once again, however, there is evidence for nonrandom abortion of seeds within fruits^{3,7}. Genetic load due to recessive lethals also implies that the aborting embryos could never survive. In fact, aborting embryos within fruits, at least in some species, can be rescued if the dominant embryo is killed¹⁰.

Charlesworth does not discriminate between embryo and fruit abortion, both of which contribute to the lowered maturation of ovules (and hence lead to the question of why so many ovules). Such distinction has important implications because the pattern of abortion of embryos within a fruit⁶ and that of fruits within a plant^{11,12} may be distinctly different. If embryo abortion is due to the genetic load, the retained fruits are expected to exhibit a normal distribution of seeds per fruit because the aborting embryos should be randomly distributed in all fruits. This however is not true and the seeds per fruit show varied but species-specific distribution patterns⁶. At least some of these patterns (for example, negatively skewed distribution of seeds within fruits) are not due to the embryo abortion; rather they are due to the seed-number-dependent selective abortion of pods^{8,11,13}, which may or may not be dependent on genetic complement of the seeds¹².

We have recently reviewed factors responsible for ovule abortion within developing fruits. Parent-offspring conflict over resource allocation and sibling rivalry can both lead to abortion of developing embryos⁶. The production of extra ovules may be a means of selecting embryos in the face of diversity of pollen genotypes reaching the stigmas⁴. Pollen competition has also been assumed to play a role in ovule abortion^{3,5,7}. These hypotheses do explain the frequently observed correlations between the levels of abortion and mating systems. Unlike the

genetic load hypothesis, these hypotheses also explain nonrandom abortions, correlations between patterns of abortions and seed dispersal and the low level of abortions in highly outcrossed taxa such as tree species in the Mimosoideae. But they do not explain postdevelopmental abnormalities in seed and seedlings observed by Wiens *et al.*; note, however, that evidence for such abnormalities is based on a sample size of only nine seeds².

There is, therefore, little evidence that segregational load is a leading factor in the abortion of zygotes and embryos in natural populations of plants. Several other hypotheses provide alternative explanations for the patterns observed in nature. These hypotheses are based on reasonable genetic models and plausible developmental, physiological and ecological mechanisms. Until more data are available, therefore, segregational load should be considered as one of the least likely possible general explanations for ovule abortions.

K. S. BAWA

Department of Biology,
University of Massachusetts,
Boston, Massachusetts 02125, USA

S. G. HEDGE

K. N. GANESHAIAH

Department of Farm Forestry,
University of Agricultural Sciences,
G K V K Campus,
Bangalore 560065, India

R. UMA SHAANKER

Department of Crop Physiology,
University of Agricultural Sciences,
G K V K Campus,
Bangalore 560065, India

CHARLESWORTH REPLIES—I completely agree that genetic load is not a very likely explanation for ovule abortion in plants. I only gave the genetic load hypothesis serious attention, because it was emphasized in the paper² on which I was commenting, and it was therefore necessary to explain the idea clearly enough to evaluate it critically. In my discussion of the hypothesis, I stated that segregational load (due to heterozygote advantage) cannot explain the data, and for mutational load that: “it seems unlikely that all this load will be expressed in the earliest stages of development so as to be detectable as reduced seed maturation. . .”, and that the lack of evidence for strong inbreeding depression argues against this interpretation for the case of *D. eurekaensis* that was under consideration.

As I also pointed out, if mutational load causes abortion “it should be random within fruits”. The evidence for nonrandom abortion in the references provided by Bawa *et al.* is relevant to this point, as is the evidence for regular differences between species that I mentioned. Both these types of evidence show that abortion does not depend entirely on the genotypes

of the aborted zygotes. This must also be true in species where it has been shown (by experimentally destroying competing embryos) that aborting embryos are not intrinsically inviable¹⁴. Such evidence also rules out maternal choice among embryos, or sibling competition, if based on the genotypes of the embryos. In many species, however, nonrandom abortion is difficult to demonstrate, and most species have some random abortion. We cannot therefore exclude a contribution of mutational load to low female fertility.

It is certainly worth considering the additional possible causes mentioned by Bawa *et al.* for the patterns of abortion and low ovule fertility seen in some plants (parent-offspring conflict, providing excess ovules so that maternal choice can occur, and pollen competition). These hypotheses are difficult to test, and are at present rather controversial, and I did not have space to mention them in my brief discussion of the question. The first is closely related to the high cost of fruit and seed maturation, discussed in my News and Views in terms of theories of resource allocation which show that this cost can lead to the evolution of low female fertility, even though potentially viable progeny are lost. Pollen competition does not seem a likely explanation for species with regularly low seed to ovule ratios, because some pollen sources should yield seeds from every ovule, which does not appear to be possible in such plants¹⁵.

DEBORAH CHARLESWORTH

Department of Ecology and Evolution,
University of Chicago,
Chicago, Illinois 60637, USA

WIENS *ET AL.* REPLY—Our letter to *Nature*² reported extensive (97.5 per cent) genetically mediated developmental failure of ovules and consequent loss of reproductive capacity in *Dedeckera eurekaensis* (Polygonaceae), a rare, palaeoendemic, monotypic genus from the Mojave desert of California.

Bawa *et al.* question both the uniqueness and interpretation of our results. They indicate that mimosoid legumes have exceptionally low fruit/flower ratios (per cent fruit set), but relatively high seed/ovule ratios (per cent seed set) giving these plants overall seed set similar to *Dedeckera* (2–3 per cent). We have confirmed these results in African *Acacia* and North American *Prosopis*. Apocynaceae and Asclepiadaceae behave similarly, as do many woody groups, and plants reproducing vegetatively. These reproductive systems, however, are not comparable to that of *Dedeckera*.

In mimosoid legumes low fruit set is genetically programmed by the maternal genome. The great majority of ovules are probably unfertilized, or if fertilized, probably abort before the activation of the zygotic genome at differentiation, or their

maturation is physiologically inhibited by earlier developing fruits. Such ovules never enter the potential seed pool.

Maternal fitness, however, is determined by the absolute number of fully viable seeds produced. In spite of the low fruit set per inflorescence, a mimosoid may produce hundreds or possibly thousands of inflorescences with one to several fruits and thousands of seeds.

In *Dedeckera* more than 95 per cent of the ovules are fertilized (flowers are uniovulate) and develop to various stages. The few filled seeds that mature occur randomly in the inflorescence; any ovule can potentially enter the seed pool. The 97.5 per cent abortion rate among *Dedeckera* ovules, however, represents only a portion of the reproductive loss. Large, vigorous plants of *Dedeckera* may produce about 64,500 flowers (and ovules). Of the 2.5 per cent of filled seeds only 3.5 per cent germinate spontaneously, and only 11 per cent are without post-developmental abnormalities. Most *Dedeckera* plants probably produce less than 50 fully viable seeds annually. Some plants are apparently totally sterile. Seed inviability of this magnitude seems unprecedented among any evolutionary successful species.

Severely reduced seed sets (inbreeding depression) are often encountered when typically outcrossing plants are selfed. Such results are of little interest. The similarly low seed sets obtained in *Dedeckera* from both outcrossing and inbreeding are precisely what make it so unusual. The genetic load in *Dedeckera* is perhaps so high that rare, recombinant gametes produced upon selfing may have as great a probability of producing a viable embryo/endosperm as outcrossed progeny. A number of phylogenetically and geographically rare plants (monotypic families) seem to have reproductive capacities similar to *Dedeckera*. These endangered plants may also exhibit unusual genetic phenomena not heretofore encountered. They require urgent study.

The failure of ovular development in general is attributable to several causes. Nonrandom abortion (maternally controlled) is exceedingly common and likely under fixed genetic control of the maternal genome¹⁶. Random ovule abortion is primarily attributable to genetic load and developmental selection^{16,17}.

Our previous letter on *Dedeckera* is essentially concerned with reduction of fitness, and presents a difficult evolutionary conundrum. If the loss of fecundity resulted in continued population decline, why would selection not reverse the trend? We proposed a heterosis (balanced-load) model that allowed for the survival of unique, highly heterozygous genotypes in an environment stressed by increasing aridity. Such genotypes should suffer drastically reduced fecundity from genetic

load. Selection, however, must first ensure individual survivorship, then fecundity. This could ultimately lead to extinction.

How can such a situation be reconciled with sibling rivalry/parent-offspring conflict hypotheses? How can they explain the high loss of seed viability, germinability, and the presence of extensive postgermination developmental abnormalities, all of which follow the release of seeds from the maternal plant? Likewise, how can they account for embryo deaths similar to those known to be controlled by developmentally lethal genes¹⁸.

These highly anthropomorphic, sociobiological hypotheses are best not applied to plants. We suggest they cannot be tested critically, and have no mechanism to explain their operation. The causal-mechanistic genetic based hypotheses, although not without difficulties, are founded on established phenomena and Occam's razor dictates their acceptance.

Since publication of our letter, we have carried out a more extensive crossing programme in which *Dedeckera* showed that of 157 self-pollinated flowers (uniovulate) one filled seed was obtained, and of 192 cross-pollinated flowers, 23 filled seeds (12.0 per cent) were obtained (N = 7 plants).

D. WIENS

E. J. KING

Department of Biology,
University of Utah,
Salt Lake City, Utah 84111, USA

D. L. NICKRENT

Department of Plant Biology,
University of Illinois,
Urbana, Illinois 61801, USA

C. L. CALVIN

Department of Biology,
Portland State University,
Portland, Oregon 97207, USA

N. L. VIVRETTE

Ransom Seed Laboratory,
P.O. Box 300,
Carpinteria, California 93013, USA

- Charlesworth, D. *Nature* **338**, 21–22 (1989).
- Wiens, D., Nickrent, N.L., Davern, C.L., Calvin, C.L. & Vivrette, N.J. *Nature* **338**, 65–67 (1989).
- Bawa, K.S. & Webb, C.J. *Am. J. Bot.* **71**, 736 (1984).
- Charnov, E.L. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2480–2484 (1979).
- Lee, T.D. *Am. Nat.* **123**, 427–432 (1984).
- Uma Shaanker, R., Ganeshiah, K.N. & Bawa, K.S. *A. Rev. Ecol. Syst.* **12**, 253–279 (1981).
- Bawa, K.S. & Buckley, D.P. *Miss. Bot. Gard. Monogr. Syst. Bot.* **29**, 243–262 (1989).
- Bawa, K.S. *Evolution* **28**, 85–92 (1974).
- Sutherland, S. & Delph, L. *Ecology* **65**, 1093 (1984).
- Ganeshiah, K.N. & Uma Shaanker, R. *Oecologia* **77**, 135–140 (1988).
- Stephenson, A.G. *A. Rev. Ecol. Syst.* **12**, 253–279 (1981).
- Stephenson, A.G. & Winsor, J.A. *Evolution* **40**, 453–458 (1986).
- Bookman, S.S. *Evolution* **38**, 1337–1349 (1984).
- Caspar, B.B. *Evolution* **31**, 495–511 (1977).
- Willson, M.F. & Price, P.W. *Evolution* **31**, 495–511 (1977).
- Wiens, D. *et al.* *Oecologia* **71**, 501–509 (1987).
- Klekowski, E.J. *Mutation, Developmental Selection, and Plant Evolution* (Columbia University Press, New York, 1988).
- Meinke, D. *et al.* in *Plant Genetics* (ed. Freeling, M.) 129–146 (Liss, New York, 1985).

Anatomy of a conflict

Shepard Krech III

Archaeological Perspectives on the Battle of the Little Bighorn. By Douglas D. Scott, Richard A. Fox, Jr, Melissa A. Connor and Dick Harmon. *University of Oklahoma Press: 1989. Pp. 309. Hbk \$24.95.*

WHAT more could possibly be said about the Battle of the Little Bighorn that has not been? An event that electrified the country the instant it happened, it continues to excite the imagination. Biographies have been written on the principals involved on both sides — George Armstrong Custer, Crazy Horse and others — as have numerous detailed accounts, from both Indian and white perspectives, of the battle itself and the historical circumstances leading up to and following this epic confrontation, on 25 June 1876, between the US Seventh Cavalry and the Sioux and Cheyenne Indians.

There are no lingering significant controversies in the interpretation of this historical record. But because there are different perspectives, and because the production and re-production of history is invariably a cultural, presentist exercise, conflicts persist over the interpretation of details in the voluminous historical archive, which probably will occupy historians well into the future. Nevertheless, when one reads on a book jacket that assembled therein is “the most compelling and convincing evidence thus far of what happened on that day in southeast Montana 113 years ago”, one can be forgiven for suspecting a publisher’s marketing department of hyperbole.

Indeed, the aims of the authors of this book are far more modest, even though at times, it seems, they also are caught up in Little Bighorn hype when they are not actively proselytizing for the revisionist potential of historical archaeology. Grasping the opportunity presented by a grass fire that removed vegetation on the Custer Battlefield National Monument, the authors (of the four, one is a military archaeologist, another an expert on firearms) undertook a comprehensive programme in historical archaeology to determine where 28 soldiers killed in a cul de sac were interred; whether marble battle-

field markers actually were placed over human remains; what the bones of troops reveal about diet, trauma and death; how combatants moved about the battleground; and what firearms were used.

The results of archaeological investigations — metal-detector surveys, shovel and backhoe excavations — are mixed. The authors failed to discover the remains

ments — “the number of bullets impacted in the defence perimeter strongly suggests that the incoming fire must have been heavy” (p. 140). Although plotting distributions leads to a “further clarification” (p. 138) of the positions of the combatants or “correlates well with the historically known combatant positions” (p. 134), it unfortunately adds little to the known historical record.

The most interesting results concern the identification and analysis, on the basis of cartridge cases and bullets, of types and quantities of firearms used by Indians and whites. More than 40 types of guns were used in the battle, and the distribution of types — even of individual weapons — can be derived from signatures that guns leave on bullets and cartridge cases. Even here,



Illuminating the past — H. Charles McBarron's painting of Custer's last stand.

of the last 28 soldiers, but they think they now know, by the use of geomorphology and archaeology techniques, where the troopers cannot be and where they are likely to be found. As for the markers, some are improperly placed, not surprisingly because the dead were ‘buried’ three or four times: the day after the débâcle, the bodies were covered; in 1877 and 1879, exposed remains were reinterred; and in 1881, disinterment was followed by reburial in a mass grave. Nine years later, marble markers were placed on the battlefield, and the following year a map was prepared. Little wonder there are discrepancies. The bones show ample signs of trauma after death, to which all historians concur and about which, for cultural context, Indian sources are most revealing.

The distributions of cartridge-case, bullet and other artefactual materials on the battleground make interesting figures, despite predictable (and wooden) assess-

however, the archaeological record in part corroborates extant archival materials: a list of guns surrendered by the Sioux and Cheyenne in 1877.

Despite their evident enthusiasm for historical archaeology, the authors wisely state that they do not believe it represents “the last work” (p. 11) in the study of the Little Bighorn. Indeed, apart from often being able to corroborate the known historical record, the conclusions that one can draw from this work are limited. Similarly, where the historical record is in dispute, the illuminations provided by historical archaeology are few, with one exception: bullet and cartridge-case distributions support Indian, not white, accounts of the battle. □

Shepard Krech III is a professor of Anthropology and director of the Haffenreffer Museum, Brown University, Box 1921, Providence, Rhode Island 02912, USA; and editor of the journal Ethnohistory.

Keeping them up in the air

Thomas A. McMahon

Bird Flight Performance: A Practical Calculation Manual. By C. J. Pennycuick. Clarendon: 1989. Pp.153. £25, \$49.95.

THE subtitle says this book is a "practical calculation manual". Right away, the prospective reader wonders what that could be. Manuals are generally not much fun for off-duty reading. There are many people who might like to read an informative book about bird flight, but few who need to make calculations. The prospective reader wonders if this is a book or a computer manual, and if it's a manual, shouldn't it be someone else's job to read it?

The first chapter doesn't make the question go away. In the very first paragraph, the author launches into a discourse about how millilitres of oxygen and grams of fat are not units of energy. He reveals here, as in other places in the book, that he is on a mission to set some of his colleagues straight about the right way of doing certain things. On p.11, for example, he draws the right and wrong methods of tracing a wing to find its area. On p.8 he says that some collections of data in the literature have to be treated with caution "because ornithologists have proved surprisingly creative in inventing wrong ways to make even the simplest measurements".

Not until the reader is well into the book does it dawn, like a bright light switching on, that this volume — manual or not — represents a beautiful and original achievement. From then on, it doesn't matter whether it is a book, a manual or a computer program (actually, it is all three). What matters is that Pennycuick has invented something here; a way of exploring a complex biological subject through first-hand revelations using a calculation model.

It is a very successful invention. Books are good tools for self-education, but Pennycuick obviously believes there are better ways to learn. The best way is not to be told the answer but to be given the opportunity to discover it for oneself. Pennycuick's invention is not at all like computer-aided instruction, where the program asks the user a question and then scolds and corrects when it thinks the answer is wrong. Using the programs provided here, the user is responsible for formulating a scientific question. An example (p.88) might be: why do birds of intermediate size use a mode of intermittently powered flight (bounding) in which the wings are folded against the body during the non-flapping phase? The user



On the wing — recently published, *The Birds of South America* by R. S. Ridgely and G. Tudor, is the first volume in an intended series of four. The initial volume is subtitled *The Oscine Passerines* and encompasses more than 700 species of songbird. Thirty-one colour plates, drawn by Guy Tudor, illustrate the text of Robert Ridgely. This work is seen as a valuable contribution to the recording and appreciation of the birds of South America — especially appropriate at a time when rapid development is devastating millions of acres of tropical habitat in South America. The subsequent volumes of *The Birds of South America* will be: *The Suboscine Passerines*; *The Nonpasserines (Landbirds)* and *The Nonpasserines (Waterbirds)*. This illustration is taken from the cover of the first volume, published by Oxford University Press, price £45.

can work out answers to such questions by obtaining measurements, not on animals, but from a calculated simulation of what the animal would do in the presumed situation. Most of the book concerns explanations of how the programs arrive at their answers. Pennycuick provides brief discussions of the physical and biological principles behind the calculations, emphasizing throughout the importance of selecting and configuring the program for the task at hand and staying within the limitations inherent in the program's assumptions. He has made it straightforward to change the assumptions when necessary, as, for example, when the air density is lower because the bird is flying at altitude.

The book is full of fresh insights I have never seen published. I liked the very simple and clear discussion of Reynolds number on p.43, where the author points out that the ratio of inertial force to viscous force acting on the air in a little cube flowing around an obstruction grows larger if you scale the whole drawing up because the ratio of the cube's mass to its surface area increases. On p.50 I was told that a snow goose body in a wind tunnel has 15 per cent less drag when the feathers are coated by hair spray. In a living bird, the feathers can be controlled by pteromotor muscles, presumably giving some of the same effect. On p.62, there is a diagram of the concertina wake, in which

the bird flexes and shortens its wings on the upstroke while flapping, so that no transverse vortices are shed, and consequently less energy is required than would be the case for flapping wings of fixed span.

A particularly nice use of the programs included with the book is discussed on p.107, where the author asks why small birds don't take advantage of the energy in the atmosphere and soar as large birds do. He shows how his programs can produce a telling calculation comparing a 5.5-kilogram crane to a 17.4-gram ovenbird, assuming both birds have 1 metre per second thermals to soar in. Whereas the crane requires only 15 per cent as much energy per kilometre to migrate using thermals as to flap along in still air, the little ovenbird uses 6 per cent *more* energy per kilometre riding the thermals. The reason the small bird comes out so badly is that its basal metabolism is a much larger fraction of the total energy requirement. Because the small bird spends almost three times as long using thermals as it would flying direct, the extra time results in an extra cost for basal metabolism that overwhelms the energy savings that would be available from gliding instead of flapping. □

Thomas A. McMahon is in the Division of Applied Sciences at Harvard University, Cambridge, Massachusetts 02138, USA.

Moving with the times

Richard Taylor

Vogel's Textbook of Practical Organic Chemistry, 5th edn. Revised by B. S. Furniss, A. J. Hannaford, P. W. G. Smith and A. R. Tatchell. Longman: 1989. Pp. 1,514. £36, \$84.95.

FORTY-one years ago, Arthur Vogel, former head of the Chemistry Department at Woolwich Polytechnic, London, published *A Textbook of Practical Organic Chemistry Including Qualitative Organic Analysis* to provide a comprehensive, single-volume reference book for organic chemists at all stages of their careers. Vogel produced revised editions in 1951 and 1956, and a fourth edition by the present team, from the same institution as Vogel (now called Thames Polytechnic), appeared in 1978.

It is fascinating to compare the contents of the five editions and to chart the changes in organic chemistry over the past four decades. The third edition included, for the first time, accounts of "electronic mechanisms in outline", syntheses with organolithium reagents, reductions with potassium borohydride, some applications of chromatographic absorption and a brief introduction to infrared and ultraviolet spectroscopy. The fourth and fifth editions, by contrast, contain large sections on chromatography, spectroscopy (including nuclear magnetic resonance and mass spectrometry), new synthetic reagents and safety, and the preparative chapters are written from a mechanistic viewpoint. The major casualty in this process of revision, the chapter on the theory of general technique, was published independently as *Vogel's Elementary Practical Organic Chemistry* (Longman, 3rd edn, 1980). The main change in the fifth edition is the addition of an introductory chapter on the design of organic syntheses, which includes discussions of selectivity, protection and stereocontrol as well as of retrosynthetic analysis and the disconnection/synthon approach. These ideas are then integrated in the preparative chapters. There is even an appendix on common synthons and their reagent equivalents — how times have changed!

A more detailed inspection of the new edition reveals the extent of the upgrading: there are new sections on many 'recent' techniques, including reactions involving air-sensitive compounds, Kugelrohr distillations, high-performance liquid and flash chromatography, and centrifugal chromatography using the Chromatotron and Kugelrohr distillations. The chapter on solvents and reagents has been

expanded to include organoborane reagents, organolithium reagents (although I was surprised that the Gilman double-titration estimation method was not covered) and many other modern synthetic reagents. The preparative chapters also reflect this modernization, with the inclusion of illustrative asymmetrical syntheses, protection reactions and transformations involving enzymes, supported reagents, phase-transfer catalysts, strong bases, and air- and moisture-sensitive reagents. About 100 new experimental procedures have been introduced in this edition, although for the first time they have not been checked in the editors'

laboratory, and the pruning of less relevant procedures and old-fashioned terms has continued (although rectified spirits still persists).

When asked a practical question, many organic chemists (including myself) invariably respond "Consult Vogel". That is the way it stays, except that one should now add "the fifth edition, of course". □

Richard Taylor is in the School of Chemical Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

■ Recently published by Longman is the 5th edn of *Vogel's Textbook of Quantitative Chemical Analysis*, revised by G. H. Jeffrey et al. Price is £34, \$87.95.

Rules of conduct

John Tyler Bonner

Dynamic Structures in Biology. Edited by Brian Goodwin, Atsuhiko Shibata and Gerald Webster. Edinburgh University Press: 1989. Pp.259. £35.

I CANNOT remember when I have worked so hard reading a book. Probably I would have put it down after the first few pages had I not agreed to review it, and had I not wanted to come to grips with 'structuralism' and find out what it has to contribute to biology. The difficulty is that if I try to explain the approach in the terms of the more dedicated of the contributors to this volume, it will not immediately seem clear and crisp. That is partly because structuralism has its origins in the writings of Levi-Strauss and Piaget, neither of whom do I find paragons of straightforward thoughts or uncomplicated prose. So I will not even attempt to give an approved definition of structuralism, with its object of revealing relationships and finding laws and principles of organization that are to become the basis of our understanding of development and evolution.

Let me begin by giving my unexpurgated version of structuralism in biology, with emphasis on how it came into being, because that clears up part of the mystery for me. For well over a hundred years, many biologists have been genuinely unhappy with evolution by natural selection, or darwinism. They feel that it is far too simplistic to say that the grandeur of evolution can be explained by a somewhat haphazard genetic variation followed by selective reproduction success. The same problem arises with the magic of an egg turning into a tree or a giraffe; how is it possible that genes can give off the kind of information needed for those wonderful developmental transformations? It is as though one were having the great mysteries of the world explained by Beatrix Potter. According to biological structuralists, evolution and development deserve

more than the ridiculously shallow answers provided by natural selection and developmental genetics. To ram the point home, a true structuralist considers them both 'reductionist', and manages to say so in such a way as to equate this with original sin.

For many, especially in earlier times, the problem was easily solved by saying that all organisms were created by God. This fully accounted for all the magnificence of evolution and the baroque yet perfect functioning and construction of higher animals and plants. But this explanation requires the right kind of faith, and as that waned other antidotes to reductionist explanations took over. There was a succession of solutions to the problem. Vitalism flourished in various forms — the *élan vital* of Bergson, the entelechies of Driesch — but such ideas are no longer socially acceptable. Evolutionists tried to circumvent the problem by either pushing for a strong form of lamarckism or fostering the idea that there were internally driven trends in evolution which they called orthogenesis. The latter is little more than camouflaged vitalism; and although lamarckism removed some of the problems, it eventually foundered because it was testable and shown to be wrong. There are many more ins and outs in this long tale of unease about darwinism, and the problem continues to be with us today. Many distinguished mathematicians and theoretical physicists reject it because they see their world as a kind of perfection that they are trying to reveal — there is no place for the messy and trivial process of natural selection.

It is clear that structuralism in biology is another manifestation of the unease generated in so many good minds by darwinism. Here is my version of how the structuralists have responded. They say that there is an inner structure to organisms that obeys laws and principles yet to be determined, and that when they are fully understood we will have a satisfying explanation for both development and evolution. In evolution there is a *Bauplan* which can change according to these

presumed rules, and that accounts for evolution. It is not clear how a primaevial bacterium can have the incipient structure of both a mammal and an angiosperm, but I assume that these laws of structure allow for — even require — a rigid, prescribed sequence of change. It is also possible, as some of the authors suggest in this symposium volume, that one might have a mixture of natural selection and these laws of structural transformation, the latter dealing with the important changes in form. It is argued, for instance by M. Ito, that convergence in evolution is not explained by the standard idea of similar selection pressures producing similar structures, but that the reverse is true: the primordial structure is the same and that is why different kinds of organisms produce the same shapes. There is even the interesting suggestion by K. Ikeda that these structural rules might be exerting their influence directly on the DNA of the genome in the form of super control elements.

The problem of development receives considerable attention from a number of authors, especially Brian Goodwin. He seeks these laws of structure and asks whether perhaps some of the reaction-diffusion models of development inspired by Alan Turing, or the mechanochemical model of Murry, Odell, Oster and others, might lead us into the structural laws. The arguments for this and other more directly genetic influences on development are intelligent and sophisticated, but Goodwin has not made me a convert. For me, the quest of the grail of developmental biology is not some hidden set of laws, but a simple unravelling of the web of mechanisms that control development. I see the experimentalist who does developmental genetics as being on the right track, and there is much in current progress to encourage that view.

There is another aspect to structure and evolution development that should be mentioned. Many authors have recently pointed out that organisms can only build in one generation, or change over many, if they use the building blocks which are available to them. They have called these 'constraints' and made much of the obvious point that neither evolution nor development can make wildly abrupt morphological shifts. It is in these very constraints where structuralists see the hidden power that will explain all.

To return to the volume, let me say that it is a mixed bag of essays. Gerry Webster introduces the topic and many of the papers directly address the idea of structuralism, some clearly also embracing darwinism. A few are wonderfully impenetrable, while others, such as the short paper of K. Ikeda, are models of clarity.

D'Arcy Thompson's *On Growth and Form* is often referred to in the book, and this adds yet another facet to the psychological problem. For Thompson a satisfying

explanation was a mathematical description. I understand and respect this, but it does not satisfy me. For the structuralist there is an unknown set of rules that govern form, but that does not satisfy me either, for I feel they may not exist and certainly it is unlikely that we will ever find out what they are. I am quite happy with vulgar natural selection and developmental genetics. □

John Tyler Bonner is in the Department of Biology, Princeton University, Princeton, New Jersey 08544, USA.

New development

Tim Hunt

Genes and Embryos. Edited by D. M. Glover and B. D. Hames. IRL/OUP Press: 1989. Pp. 228. Hbk. £27, \$59; pbk. £18, \$38.

UNDERSTANDING how genes control embryonic development is one of the most important and exciting aims of biology. Thomas Hunt Morgan and his colleagues at Columbia University understood this when they provided the basis of modern genetics, but they made only modest progress in discovering how genes actually shape the organism. In the fruitfly *Drosophila melanogaster*, however, they chose the right organism to solve their problems, as Morgan's followers have since triumphantly demonstrated. The many advances in cell and molecular biology of the past 20 years or so were required, however, to confirm and flesh out the *Theory of the Gene* (Yale University Press, 1926).

But let us not forget that creatures that 'don't have genetics' have also shed light on development: amphibians led and arguably continue to lead the way in the study of induction (with nematodes coming up fast on the inside). Good progress (considering the formidable technical problems posed by development within a womb) has been made in describing and analysing mouse embryogenesis. Surprisingly, many of the key genes that control development in flies are found in vertebrates too, where they probably work in not-too-dissimilar ways. In fact, a tremendous amount is now known, and astonishing principles are emerging at an accelerating pace.

Now is a good moment to tell the world about these new understandings. People like me sense that these are stirring times but have trouble keeping their segments, parasegments, gap- and pair-rule genes in their correct conceptual slots. Stripey flies, mammalian homoeobox genes and growth-factor receptor homologues in worms and flies are frequently reviewed in specialist journals, but it would be wonderful if someone could put it all together

and tell, if not *the* story, at least a well-illustrated outline of what the story is going to be. This will take more than a minireview in *Cell* or an article in *Trends in Genetics*. Perhaps the nearest approach to this ideal is chapter 16, "Cellular Mechanisms of Development", in *Molecular Biology of the Cell* (eds B. Alberts *et al.*, 2nd edn; Garland Press, 1989).

I had hoped that *Genes and Embryos* was the book I was looking for, but sadly it is not. Perhaps its opening sentence gives the game away: "This book reviews the exciting discoveries made over recent years which contribute to our understanding of development at a molecular level." It is the *molecular* that presents the problem, I think, for the authors are all excellent and their organisms apt.

Almost half the book is allotted to *Drosophila*; Kathryn Anderson, Michael Levine and Katherine Harding give commendably up-to-date accounts of developments in this fast-moving and tremendously impressive field. But their densely worded chapters are too sparsely illustrated for easy comprehension. My favourite chapter is that by Kenneth Kemphues on the nematode worm. He achieves a nice balance of background information, methodological explanation, shows some striking pictures and provides food for thought. Comparatively less is understood of frogs and mice, and it is understandable that the chapters by Tom Sargent and Ian Jackson are increasingly devoted to descriptions of particular genes that happen to have been studied and, particularly for the mouse, to purely methodological considerations. Less attempt is made to give an integrated or comparative view of the principles of development of these more complicated organisms. Perhaps it is still a bit too soon.

The dream book of late twentieth-century embryology thus remains to be written. It will need expert authors who meet to argue and refine their ideas and their presentation as well as a sympathetic editor and a gifted illustrator. The result will be a fantastic story — one that I had never thought to hear in my lifetime. □

Tim Hunt is in the Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK.

In paperback

■ Just published by Cold Spring Harbor Press is *Molecular Genetics of Early Drosophila and Mouse Development*, edited by Mario R. Capecchi. Part of the series Current Communications in Molecular Biology, the book reports the proceedings of the Banbury Center meeting held on 20–23 April 1989. Price is \$24.00.

■ Also just published in the same series is *Polymerase Chain Reaction*, edited by Henry A. Erlich, Richard Gibbs and Haig H. Kazazian Jr. Price is \$22.00.

Chemical asymmetric synthesis

John M. Brown & Stephen G. Davies

Thalidomide comes in two forms: a left-handed compound which is a powerful tranquilizer, and a right-handed version which can disrupt fetal development, causing severe handicap. As a necessary consequence of synthetic methods available in the early 1960s the two forms were present in equal proportions in the manufactured drug, with catastrophic consequences. The thalidomide story is perhaps the most painful reminder of the importance of stereochemistry—the spatial ordering of groups in a molecule can be as influential as the chemical nature of the groups themselves. A principal problem in organic synthesis is, therefore, the development of methods for producing complex molecules with a stereochemically defined structure.

OVER millions of years the complex architectures of enzymes (average relative molecular mass 50,000) have developed into efficient and specific templates for only one hand of organic reactant. Since the first synthesis of an organic compound by a chemist in 1824, relatively small and simple reagents (average relative molecular mass <100) have been developed to achieve the synthetic transformations necessary for the construction of organic molecules. Despite enormous advances in synthetic methods, which have made the construction of virtually any desired compound feasible, the problems of producing single-handed compounds seemed insurmountable and synthetic compounds were always produced as equal mixtures of left and right hands. Recently, however, several breakthroughs in synthesis have occurred whereby more sophisticated, single-handed reagents or catalysts (relative molecular mass <800) have been invented which control the handedness of the product compounds with efficiencies rivalling those of enzymes.

In structures of biological interest, such as natural products or synthetic enzyme inhibitors, there are one or more (often several) saturated carbon atoms with four different substituents attached and this tetrahedral arrangement can be either of the possible hands. Figure 1 shows the hydration of but-1-ene in which the addition creates such a chiral centre and here, because no selectivity has been imposed, both enantiomers are formed. A chemical reagent that is itself chiral may select in favour of one of these, producing a homochiral product. If the reactant already possesses a chiral centre then the products, diastereomers, are no longer simply related as object and mirror image and they have different chemical and physical properties. This is illustrated in Fig. 1a by the hydration of homochiral 3-hydroxybut-1-ene. The art of asymmetric synthesis is in controlling reactions so that one enantiomer or diastereomer is produced predominantly or exclusively¹⁻³.

Classical methods for the production of homochiral compounds generally involve some form of resolution experiment, such as Pasteur's manual separation of enantiomorphous crystals or the fractional crystallization of diastereomers⁴. If the asymmetric centre is to be introduced by synthesis then one of several different techniques must be applied:

- The reagent or reactant may contain a chiral auxiliary, a group in the vicinity of the reaction site that controls the stereochemistry and which is easily removed afterwards, or
- A homochiral catalyst may be used, which can be an enzyme or a synthetic product, or
- A readily available homochiral molecule, often from a natural source, may be incorporated as a building block—the 'chiral pool' approach.

Examples of these are shown in Fig. 2. The first two are examples of true asymmetric synthesis because chirality is created in the course of reaction, and we will concentrate on such approaches here.

The advances in chemical methods of asymmetric synthesis over the past five years⁵⁻⁷ have been such that the enantiomer excess (% desired – % undesired enantiomer) achieved in a good asymmetric synthesis can be >99%, and 95% is fairly routine. Such successful asymmetric syntheses require that information is transmitted efficiently from a pre-existing homochiral source to the reaction site. This molecular recognition process involves a combination of attractive forces (electrostatic or van der Waals) and steric repulsion.

A high proportion of recent asymmetric syntheses use organometallic or coordination chemistry. There are several reasons for this—in particular the ability of metals to complex with the common functional groups in organic molecules in a reversible manner, and the ability to activate simple reactants such as hydrogen and carbon monoxide. Also, the reactive complexes are well defined spatially, which is vital for stereochemical control. Physicochemical techniques that can elucidate three-dimensional structure are therefore important in developing our understanding of asymmetric synthesis, which is in essence a phenomenon of chemistry in the third dimension. Solution spectroscopic techniques, especially nuclear magnetic resonance, are used to some extent. X-ray crystallography, however, is still the most important tool, even though asymmetric reactions normally take place in the liquid phase so that an extrapolation from the solid structure is required.

Homochiral reagents and auxiliaries

Chirality does not occur spontaneously, it must be imposed during synthesis. In general this is achieved using a tightly

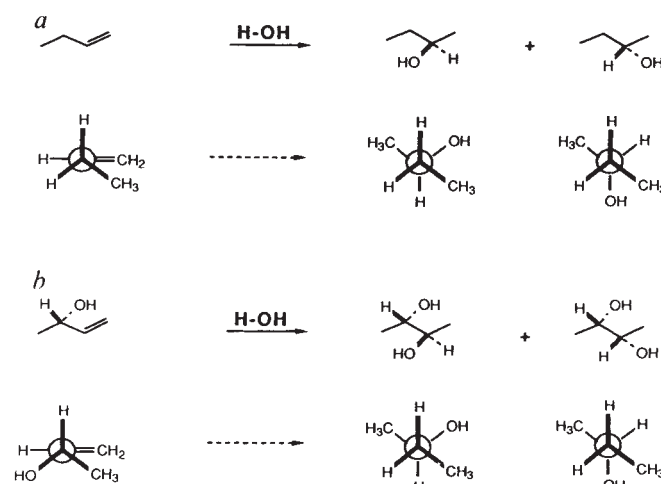


FIG. 1 Stereochemistry of hydration of *a*, but-1-ene and *b*, homochiral 3-hydroxybut-1-ene.

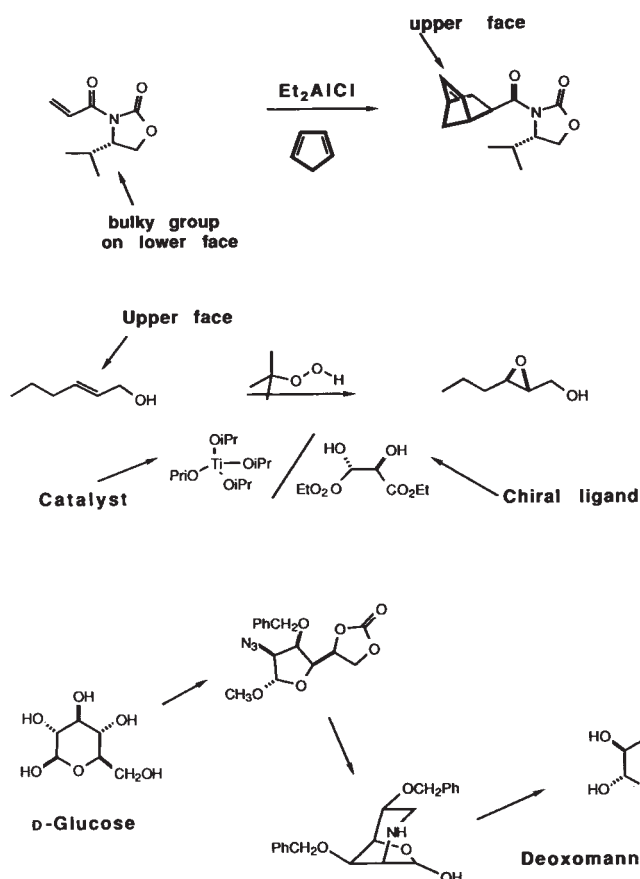


FIG. 2 Examples of asymmetric synthesis. a, The use of a chiral auxiliary; b, The use of a chiral catalyst; c, The 'chiral pool' approach. Abbreviations for the functional groups: Et, ethyl, C_2H_5 ; iPr, isopropyl, C_3H_7 ; Ph, phenyl, C_6H_5 .

controlled homochiral environment. The homochiral environment is generally provided by the attachment of a chiral appendage to one of the reactants (substrate or reagent) at the stage of the synthesis in which the new chiral centres are generated². The most general approach involves the attachment of a chiral fragment to the substrate. The three-stage approach of prior attachment of the fragment, asymmetric synthesis and disconnection of the new chiral fragment makes up the chiral-auxiliary approach whereas catalytic asymmetric synthesis is achieved when these three stages occur spontaneously under the reaction conditions.

Homochiral auxiliaries. The requirements of any homochiral auxiliary are that it must be easy to introduce into the substrate, it must be able to transfer its chirality efficiently during the synthetic transformations and finally it must be readily removed without damage to the substrate. Here we describe how one such chiral auxiliary, which we have developed (refs 8–12) at Oxford, meets these requirements together with its use in the asymmetric synthesis of homochiral (–)-captopril, a potent

inhibitor of angiotensin converting enzyme and one of the top-selling drugs worldwide.

The iron complex (cyclopentadienyl)iron(carbon monoxide)-(triphenylphosphine)(acetyl) $[(C_5H_5)Fe(CO)(P(C_6H_5)_3)COCH_3]$ is readily prepared, and commercially available as either enantiomer. It comprises the chiral auxiliary $[(C_5H_5)Fe(CO)(P(C_6H_5)_3)]$ attached to the two-carbon acetyl fragment $[COCH_3]$ (refs 8–10). This acetyl fragment forms the basis for synthesis of virtually any target molecule, for example captopril¹¹. The X-ray crystal structure of $[(C_5H_5)Fe(CO)(P(C_6H_5)_3)COCH_3]$ is shown in Fig. 3. The geometry about the central iron atom is close to octahedral—three of the six coordination sites of the iron are occupied by the cyclopentadienyl ligand with the carbon monoxide, triphenylphosphine and acetyl groups occupying the remaining three. The octahedral geometry requires the three groups carbon monoxide, triphenylphosphine and acetyl to be mutually orthogonal. The relative sizes of the three groups which comprise the auxiliary are triphenylphosphine (very large), cyclopentadienyl (medium) and carbon

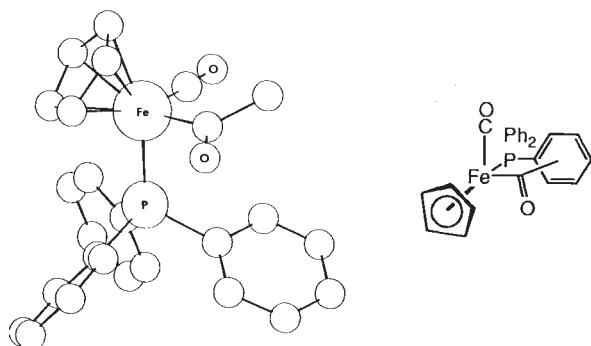


FIG. 3 X-ray crystal structure of the homochiral iron acyl complex.

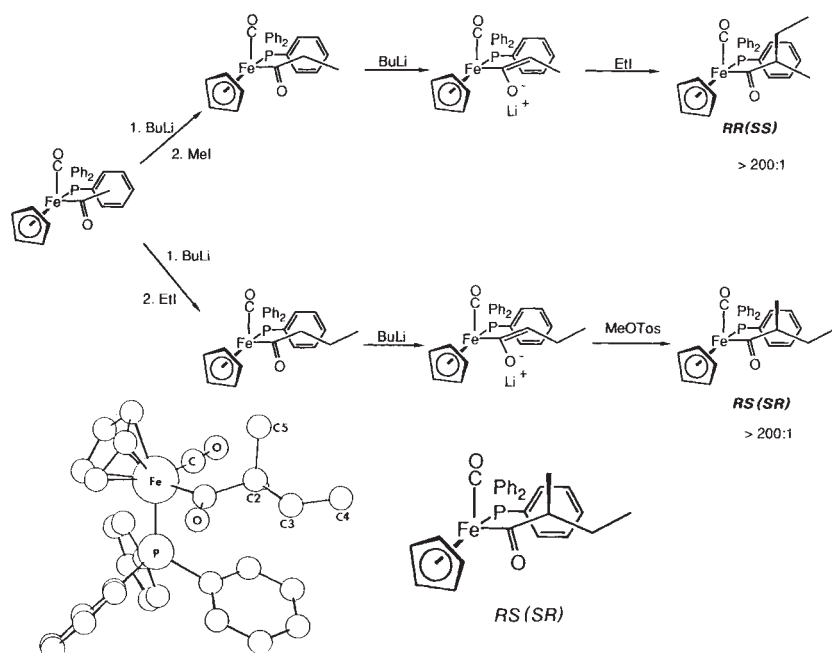


FIG. 4 The alkylation pathways for iron enolate complexes; an X-ray crystal structure of the lower product indicating the stereochemistry at C₂ is shown. Bu, butyl is C₄H₉; Tos, *p*-toluenesulphonyl is C₇H₇SO₂.

monoxide (very small). This difference in bulk of the ligands in the auxiliary orientates the attached acetyl ligand so as to minimize steric interactions. The preferred orientation places the small (relative to methyl) oxygen atom of the acetyl anti to the carbon monoxide ligand, that is, between the large and medium groups (Fig. 3). The triphenyl phosphine ligand arranges its phenyls, again to minimize mutual steric interactions, in the form of a rotor. Although this rotor is freely rotating, one blade (phenyl group) shields, at all times, one face of the acetyl group (Fig. 3) whereas the other face is essentially unencumbered. The acetyl group is thus in a fixed orientation in a chiral environment ready for reagents and reactants to approach only from the unhindered face. These ingredients are all that are required to impose stereocontrol and achieve asymmetric synthesis.

Protons on carbons adjacent to carbonyl groups are acidic^{9,12}, so removal of a proton from the acetyl, to generate an enolate, is readily achieved with strong base (butyl lithium). Trapping this enolate, which is nucleophilic at the carbon, with the electrophile ethyl iodide forms a new carbon-carbon bond. Overall the acetyl group has been converted to the four-carbon butanoyl fragment (Fig. 4). The butanoyl group prefers to orientate itself

with the terminal two-carbon fragment as far away from the very bulky auxiliary as possible; this exposes only one of the remaining two acidic protons to base. Removal of an acidic proton from the butanoyl group generates the corresponding enolate with the ethyl group *cis* to the enolate oxygen. Trapping this enolate with the electrophile methyl iodide forms a new carbon-carbon bond and generates the 2-methylbutanoyl complex (Fig. 4). The addition of the methyl group can only occur from the unhindered face of the enolate thus giving rise to the stereocontrolled formation of a new chiral centre. Figure 4 also shows the X-ray crystal structure of the product 2-methylbutanoyl complex. The chiral auxiliary maintains the same structural features as before. Looking at this structure, it should be easy to envisage the controlled construction of the chiral centre. The final step involved the formation of the C₂-C₅ bond with C₅ approaching C₂ away from the triphenylphosphine and C₃ and C₄ away from the iron. Overall two of the hydrogens of the original acetyl group have been replaced sequentially by different organic fragments. The method is very versatile, the stereocontrol is very good and the enantiomer produced for a given enantiomer of the chiral auxiliary depends only on the

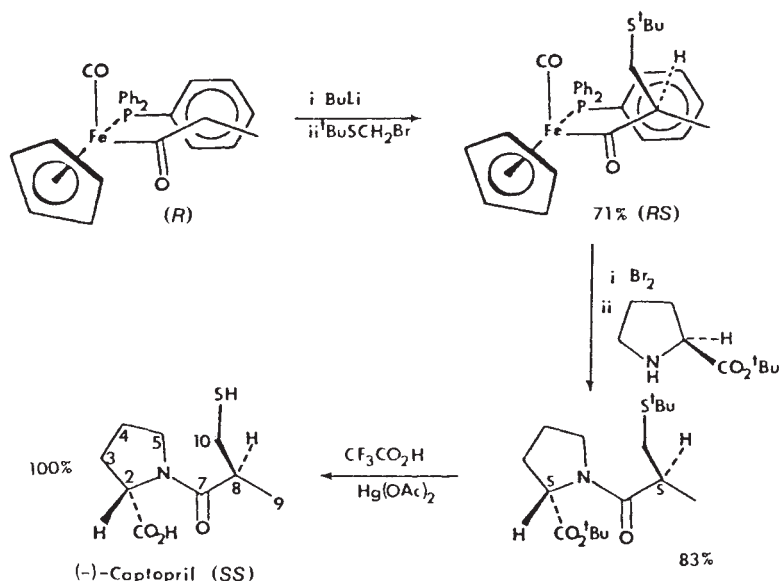


FIG. 5 The iron acyl route to SS-captopril. Ac, acetyl is COCH₃.

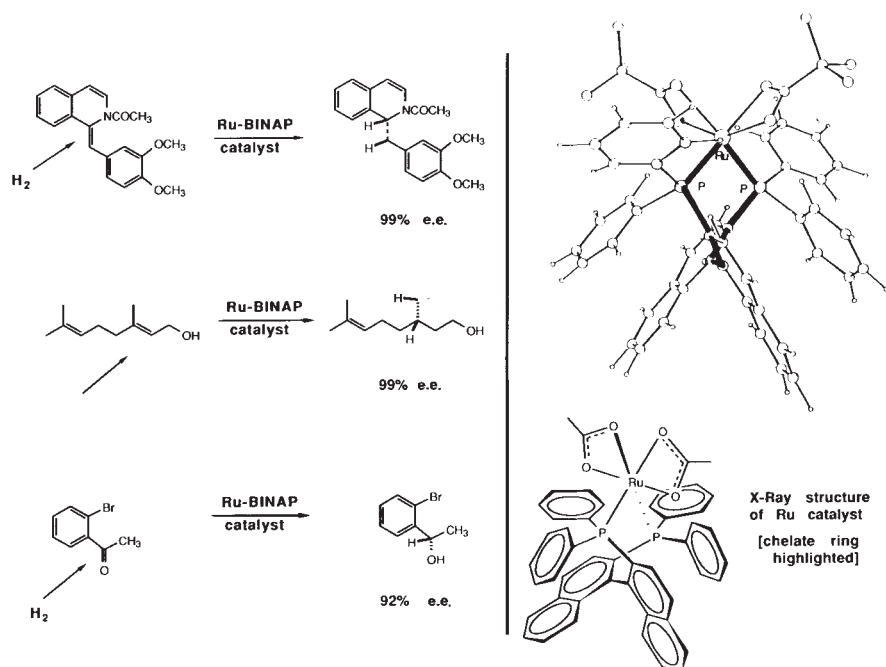


FIG. 6 Ruthenium asymmetric hydrogenation with the BINAP complex shown. e.e., enantiomer excess.

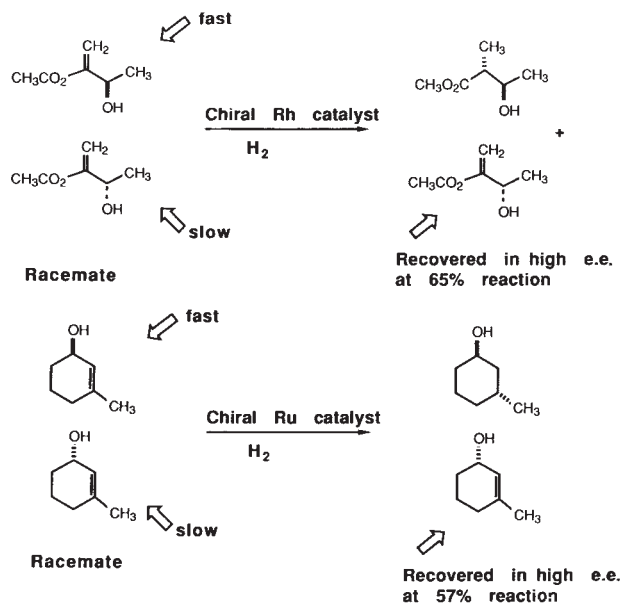
sequence of the carbon-carbon bond forming steps.

Removal of the chiral auxiliary is readily achieved by one-electron oxidants such as bromine. Oxidation in the presence of water, alcohols or amines leads to the formation of carboxylic acids, esters and amides respectively. The stereochemical integrity of the new chiral centre is not compromised in the release from the chiral auxiliary⁸⁻¹⁰.

The asymmetric synthesis of (-)-captopril is illustrated in Fig. 5. Starting from the parent iron acetyl complex the required chiral centre is constructed as before by the sequential introduction of a methyl and a protected thio hydroxymethyl group. Removal of the auxiliary in the presence of the *t*-butyl ester of L-proline forms the required amide bond to yield doubly protected captopril. Simple deprotection releases (-)-captopril that is identical in all respect to an authentic sample.

Homochiral catalysts

Enzymes in synthesis. Although we focus on chemical rather than biochemical methods, we consider the latter briefly here.



Structure of Ir complex of reactant analogue

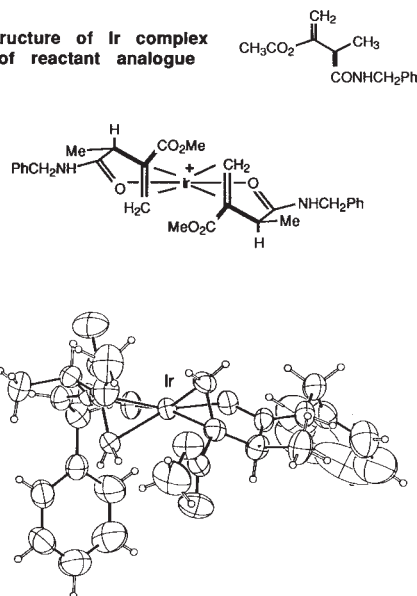


FIG. 7 Catalytic kinetic resolution in hydrogenation with rhodium and ruthenium complexes.

the reaction in organic solvents²¹. The development of homogeneous hydrogenation was a principal factor in the award of the 1973 Nobel prize to Wilkinson. Many groups realized the potential of this reaction in asymmetric synthesis. By 1977 it had been discovered that the double bond of dehydroamino acids could be hydrogenated to one enantiomer of the amino acid with essentially complete specificity. The critical insight, for which major credit must be given to H. B. Kagan and W. S. Knowles, was the possibility of using chelating biphosphines as chiral ligands^{19,20}. The recognition process involves coordination of the dehydroamino acid (introduced as an amide) to rhodium through both the olefin and amide group. The presence of a secondary binding site is a familiar concept to enzymologists and it proves to be a critical feature for asymmetric synthesis. Thus rhodium asymmetric hydrogenation is not a general reaction—substrates that lack secondary binding are not reduced with high enantioselectivity. Nevertheless, this reaction remains one of the best chemical routes to optically pure amino acids of D- or L-configuration and was used in an industrial-scale synthesis of L-DOPA (for clinical treatment of Parkinson's disease) for several years. Many homochiral ligands are effective provided that they can form chelates, and they probably operate by a common mechanism²⁰.

The next major breakthrough came in 1986–7. In early work, Wilkinson, James and others had shown that ruthenium complexes were effective for the catalytic hydrogenation of simple olefins²¹. The field was not developed to the same extent as rhodium catalysis until the group led by R. Noyori prepared ruthenium complexes of the homochiral ligand BINAP (most ligands for asymmetric catalysis have unwieldy chemical names so acronyms are used). These have proved to give extraordinary specificity for a broad range of asymmetric reductions^{22,23}. The reaction is not confined to olefins, for many unsymmetrical ketones can be hydrogenated to homochiral alcohols. Some typical reaction sequences are illustrated in Fig. 6, along with a crystal structure of the catalyst precursor, a biphosphine ruthenium bis-acetate. The reason for its efficacy is not yet properly understood, but the ligand can be viewed as a chiral scaffold, with the P-aryl rings rather rigidly defined in space, occupying two coordination sites of hexacoordinated ruthenium. In the catalytic cycle the acetates are displaced so that four coordination sites are made available. Three sites are taken up by two adjacent hydrides and by the auxiliary binding group. The remaining site is available for the olefin or ketone such that only one of its two faces can be accommodated, leading to the observed specificity.

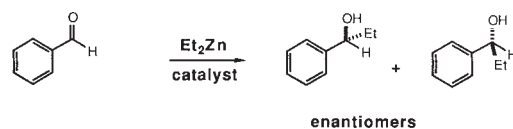
Catalytic kinetic resolution²⁴. The examples of asymmetric hydrogenation given above involved substrates that have an auxiliary binding group in the vicinity of the reduction site. If this is itself at an asymmetric centre, then further stereochemical control can be exercised. With any catalyst, chiral or otherwise, the new asymmetric centre may bear a defined stereochemical relationship to the existing one—the binding group may direct the course of reduction. The cationic rhodium complexes that are effective for asymmetric hydrogenation also work well in such directed hydrogenations, with better than 99% of one diastereomer produced^{25,26}. Further, the substrate tolerance is much greater than in the former reaction, so that hydroxy, carboxy and acylamino are all efficient directing groups.

The course of these directed hydrogenations is controlled by steric effects that result because of the binding to rhodium. The result can be predicted from molecular models which clearly show that coordination of one face of the olefin is strongly unfavourable on steric grounds. The idea is reinforced by an X-ray structure of a typical reactant coordinated to iridium—it has the same relative configuration as the product of directed hydrogenation (Fig. 7) (N. W. Alcock, J. M. B., A. Conn, A. P. James and R. J. Taylor, manuscript in preparation).

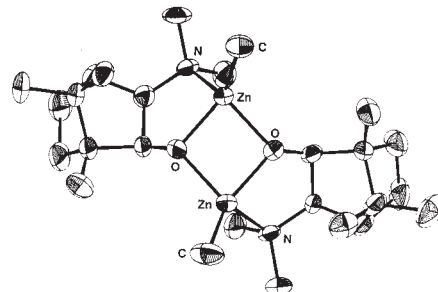
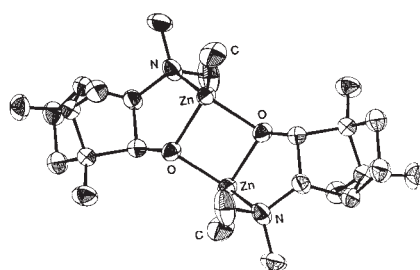
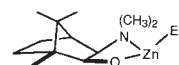
If the chiral starting material is racemic (equal amounts of both hands) and the catalyst homochiral then kinetic resolution may occur. This requires that the two hands of substrate react at sufficiently different rates. In this case the fast-reacting enantiomer is increasingly depleted as reaction proceeds, and the starting material tends towards optical purity. The relative rate of fast and slow-reacting enantiomers is expressed as a selectivity factor *S*; for an *S* value of 50 an enantiomer excess of 95% is obtained at 55% reaction (45% recovery of optically active starting material) whereas an *S* value of 10 leads to an enantiomer excess of 95% at 65% reaction (35% recovery of starting material). For the reductions catalysed by rhodium complexes, shown in Fig. 7, *S* values of between 10 and 25 have been recorded, making this an effective route to optically enriched derivatives of acrylic acid, a useful building block in synthesis.

The ruthenium complexes discussed earlier have similar efficiency in kinetic resolution of acylates, but require hydrogen pressures greater than 1 atmosphere. They can also be used for the kinetic resolution of some simple cycloalkenols (see Fig. 7) where the rhodium catalysts are ineffective.

Catalytic amplification of chirality. The reaction of organometallic reagents with aldehydes is one of the most common methods for the formation of a new carbon-carbon bond. It was discovered recently that organozinc reagents such as Et₂Zn will



Catalyst is monomeric zinc complex



X-ray structures of complexed methylzinc dimers

FIG. 8 Aminoalcohol-catalysed addition of diethylzinc to aldehydes.

add to aldehydes in the presence of a catalytic amount of a β -aminoalcohol²⁷⁻³⁰. Further, if the catalyst is homochiral, the resulting secondary alcohol is formed with a high enantiomer excess. This proves to be a valuable procedure, especially for aromatic products. The catalysts can be derived from the natural products camphor, proline or ephedrine by conventional chemical synthesis.

The catalyst for this reaction forms a complex with the organozinc reagent that has been isolated in crystalline form. It proves to be a dimer that is characterized by a central core unit—a four-membered ring of alternating zinc and oxygen atoms. This dimer possesses a C_2 symmetry axis. Noyori's group also prepared the corresponding racemic dimer which has a similar structure but with a plane of symmetry because it contains one molecule of each enantiomer. This has also been characterized by X-ray crystallography (Fig. 8). In solution, the binding constant for formation of the racemic dimer is very much greater than that for formation of the optically active dimer. Conversely, the racemic complex dissociates to monomer much less readily. For the catalytic addition reaction, it is the monomeric complex formed by dissociation of the dimer which is the active catalyst. Hence the optically active complex is a much more reactive catalyst than the racemic complex.

These observations have been exploited in a novel experiment in which chirality is apparently created^{31,32}. Consider a catalytic zinc complex in which the initial enantiomer excess of the β -aminoalcohol is 10%. Most of this will be present in solution as racemic dimer which is a rather unreactive catalyst. This leaves a small amount of homochiral dimer which readily dissociates to a reactive catalyst for the aldehyde addition reaction. Consequently, the optical yield is almost as high as that obtained with a homochiral catalyst. This apparent conjuring trick has no immediate practical importance but may help in the design of the next generation of asymmetric catalysts. It has an obvious relevance to the debate on the primordial origins of chirality, because it demonstrates an amplification process with a purely chemical mechanism.

The future

The differing pharmacological effects of the two enantiomers of chiral molecules are now well documented. Even if the inactive

enantiomer lacks undesirable side effects, administering the potent enantiomer alone halves the dose and alleviates stress on the patient's already compromised metabolism. We are at the watershed of asymmetric synthesis—in the near future it will be common practice to synthesize all potential new drugs as single enantiomers and there is already pressure from regulatory agencies in this direction.

Although we have concentrated on drugs in the above article many other exciting applications are unfolding, now that the synthesis of single enantiomers is becoming more general. The bulk properties of polymers constructed from single enantiomers of monomers are very different from those constructed from racemates. A requirement of new materials for molecular electronics is that they be non-centrosymmetric. This is, by definition, a property of enantiomerically pure materials. Future generations of, for example, optical data storage devices may depend on enantiomerically pure organic compounds.

The recent advances in chemical asymmetric synthesis that we have reported here need to be compared to enzyme methods, in which there have been recent developments in site-directed mutagenesis. The two approaches are complementary and will contribute in distinct ways. As a broad generalization, the methods of biotechnology are most suited to the production of natural molecules, or of closely related homochiral compounds, and spectacular selectivity can be achieved. By contrast, potentially all homochiral compounds are accessible in enantiomerically pure form by chemical asymmetric synthesis.

In many ways the chemical methods described here are still quite primitive. There is still much to be learned about the recognition processes which are at the core of asymmetric synthesis, and the trend will be towards tailoring the structure of reagents and catalysts to take account of increased understanding gained through mechanistic, structural and computer modelling experiments. The strength lies in versatility, because the only constraint on design of these reagents and catalysts is the ingenuity of the research scientist. We expect that the next decade will see advances in chemical approaches to asymmetric synthesis that will outstrip biological methods in terms of efficiency, versatility and output. □

J. M. Brown and S. G. Davies are at the Dyson Perrins Laboratory, South Parks Road, Oxford OX1 3QY, UK.

- Pratt, A. J., Davies, S. G., Brown, J. M. & Fleet, G. W. *J. Chem. Br.* **25**, 259-268 (1989).
- Morrison, J. D. (ed.) *Asymmetric Synthesis* Vol 1-5 (Wiley, New York, 1983-5).
- Brunner, H. *Topics Stereochem.* **18**, 129-248 (1988).
- Jacques, J., Willen, S. H. & Collet, A. *Enantiomers, Racemates and Resolutions* (Wiley, New York, 1981).
- Evans, D. A., Chapman, K. T. & Bisaha, J. *J. Am. chem. Soc.* **110**, 1238-1256 (1988).
- Gao, Y. et al. *J. Am. chem. Soc.* **109**, 5765-5777 (1987).
- Fleet, G. W. J., Ransden, N. G. & Witty, D. R. *Tetrahedron* **45**, 327-336 (1989).
- Davies, S. G. et al. *Phil. Trans. R. Soc. A* **326**, 619-631 (1988).
- Davies, S. G. *Pure appl. Chem.* **60**, 13-20 (1988).
- Davies, S. G. et al. *Bull. Soc. chim. Fr.* 608-630 (1987).
- Bashlades, G. & Davies, S. G. *Tetrahedron Lett.* **28**, 5563-5564 (1987).
- Brown, S. L., Davies, S. G., Foster, D. F., Seeman, J. I. & Warner, P. *Tetrahedron Lett.* **27**, 623-626 (1986).
- Porter, R. & Clark, S. (eds) *Enzymes in Organic Synthesis, Ciba Foundation Symp.* 111 (Pitman, London, 1985).
- Schneider, M. P. (ed.) *Enzymes as Catalysts in Organic Synthesis* (Riedel, Dordrecht, 1986).
- Jones, J. B. *Tetrahedron* **42**, 3351-3400 (1986).
- Yamada, H. & Shimizu, S. *Angew. Chem., Int. Edn Engl.* **27**, 622-640 (1988).
- Pratt, A. J. *Chem. Br.* **25**, 282-287 (1989).
- Bednarski, M. D. et al. *J. Am. chem. Soc.* **111**, 627-635 (1989).
- Kagan, H. B. *Comprehensive Organometallic Chemistry* Vol. 8, 560-600 (eds Wilkinson, G., Stone, F. G. A. & Abel, E. W.) (Pergamon, Oxford, 1982).
- Brown, J. M. & Chaloner, P. *Catalysis with Metal-Phosphine Complexes* Ch. 4, 131-170 (ed. Pignolet, L.) (Plenum, New York, 1984).
- James, B. R. *Comprehensive Organometallic Chemistry* Vol. 8, Ch. 51 (eds Wilkinson, G., Stone, F. G. A. & Abel, E. W.) (Pergamon, Oxford, 1982).
- Takaya, H. et al. *J. Am. chem. Soc.* **109**, 1596-1598 (1987).
- Kitamura, M. et al. *J. Am. chem. Soc.* **110**, 629-631 (1988).
- Kagan, H. B. & Fiaud, J.-C. *Topics Stereochem.* **18**, 249-330 (1988).
- Brown, J. M., Cutting, I. & James, A. P. *Bull. Soc. chim. Fr.* 212-219 (1988).
- Brown, J. M. *Angew. Chem., Int. Edn Engl.* **26**, 191-204 (1987).
- Oguni, N. & Omi, T. *Tetrahedron Lett.* **25**, 2823-2827 (1984).
- Kitamura, M., Suga, S., Kawai, K. & Noyori, R. *J. Am. chem. Soc.* **108**, 6071-6072 (1986).
- Soai, K., Okawa, A., Kaba, T. & Ogawa, K. *J. Am. chem. Soc.* **109**, 7111-7119 (1987).
- Noyori, R., Suga, S., Kawai, K., Okada, S. & Kitamura, M. *Pure appl. Chem.* **60**, 1597-1612 (1988).
- Oguni, N., Matsuda, Y. & Kaneko, T. *J. Am. chem. Soc.* **110**, 7877-7879 (1988).
- Kitamura, A., Okada, S., Suga, S. & Noyori, R. *J. Am. chem. Soc.* **111**, 4028-4036 (1989).

A 17,000-year glacio-eustatic sea level record: influence of glacial melting rates on the Younger Dryas event and deep-ocean circulation

Richard G. Fairbanks

Lamont-Doherty Geological Observatory of Columbia University, Palisades, New York 10964, USA

Coral reefs drilled offshore of Barbados provide the first continuous and detailed record of sea level change during the last deglaciation. The sea level was 121 ± 5 metres below present level during the last glacial maximum. The deglacial sea level rise was not monotonic; rather, it was marked by two intervals of rapid rise. Varying rates of melt-water discharge to the North Atlantic surface ocean dramatically affected North Atlantic deep-water production and oceanic oxygen isotope chemistry. A global oxygen isotope record for ocean water has been calculated from the Barbados sea level curve, allowing separation of the ice volume component common to all oxygen isotope records measured in deep-sea cores.

CHANGES in continental ice volume during the last deglaciation generally have been inferred from two indirect sources: (1) geological mapping of retreating ice margins, and (2) the marine oxygen isotope record. Direct estimates of glacial ice volume based on measured glacio-eustatic sea level changes are well documented from 9,000 years BP (before present)¹, but are poorly constrained before this time. Ruddiman^{2,3} recently reviewed evidence supporting three different deglaciation models: (1) the 'smooth deglaciation' model with the most rapid change centred at 11,000 years BP; (2) the 'French two-step' deglaciation model with maximum melting rates from 14,000 to 12,000 years BP and from 10,000 to 7,000 years BP separated by a mid-deglacial pause with little or no ice volume loss; and (3) a 'younger Dryas' deglaciation model, which is distinguished from the French two-step model by a mid-deglacial reversal with significant ice growth from 11,000 to 10,000 years BP. Ruddiman concluded that the available evidence from marine and terrestrial records favoured the smooth deglaciation model. He further suggested that the stepwise decreases of $\delta^{18}\text{O}$ ($=[(^{18}\text{O}/^{16}\text{O})_{\text{sample}}/(^{18}\text{O}/^{16}\text{O})_{\text{standard}}] - 1) \times 10^3$) apparent in many curves may not indicate sudden decreases in ice volume, but instead may be an artefact of a deep-ocean warming.

I sought to test the validity of these models, each with their own important climatic and oceanographic implications, by extending the existing coral-reef sea level curve¹ to the last glacial maximum. The fossil coral reefs surrounding Barbados are a very good system for examining the detail and timing of the last deglaciation. Here I show that the sea level curve derived from the Barbados reefs supports the French two-step model⁴. I then review the evidence for the oceanic and climatic response to this two-step deglaciation, with particular emphasis on understanding the cause of the Younger Dryas event.

Barbados corals

The Caribbean reef-crest coral *Acropora palmata* (Lamarck) is^{1,5} the best available sea level indicator for Pleistocene studies. This species is generally restricted to the upper five metres of water and dominates the reef-crest community in many locations. *A. palmata* is one of the fastest growing corals and is often found in an interlocking reef framework on the windward side of Caribbean islands. Its massive size minimizes post-depositional transport and compaction, and its abundance and generally pristine aragonite skeleton make sampling for ^{14}C or $^{230}\text{Th}/^{234}\text{U}$ age dating relatively easy.

To reconstruct glacio-eustatic sea levels, I chose to core *A. palmata* on the south coast of Barbados for several reasons: (1) the offshore bathymetry is gently sloping compared with the precipitous drop surrounding many Caribbean volcanic islands or the Bahama Bank; (2) Barbados is rimmed by three parallel offshore ridges⁶ for which I had extremely detailed bathymetric maps. Macintyre has suggested⁶ that these ridges are associated with past sea level changes, comprising potentially thick sequences of *A. palmata*; (3) onshore exposure of the *A. palmata* facies deposited during the penultimate deglaciation⁵ (isotope stage 6-5) indicates that the south coast of Barbados is being uplifted at $\sim 34 \text{ cm kyr}^{-1}$. Field mapping of these onshore *A. palmata* deposits indicated that throughout the history of reef growth on Barbados, the most extensive areal and vertical distribution of *A. palmata* occurred on the south and south-east coast; (4) Barbados is located in the western tropical Atlantic where relative sea level is least affected by changes in the equipotential surface that are due to the changing gravitational attraction of the diminishing Northern Hemisphere ice sheet⁷. Possible drawbacks of selecting Barbados include: (1) the assumption that the late Quaternary mean uplift of Barbados is without significant episodic uplift or subsidence; (2) the dipstick model of using oceanic islands to measure net ocean sea level changes may be too simple or may not apply to Barbados because it is situated on top of an accretionary prism between two oceanic plates. The only adequate way to test these geophysical assumptions is to perform the same drilling and dating experiment at other locations.

Barbados cores and their radiocarbon ages

Between 18 November and 6 December 1988, 16 cores were drilled from the RV *Ranger* offshore of the south coast of Barbados using a wire-line coring rig with a four-point mooring system. The purpose of this cruise was to core a continuous vertical sequence of *A. palmata* deposited during sea level rise associated with the last deglaciation. The stratigraphy of the ten longest cores and the corresponding ^{14}C ages are illustrated in Fig. 1. These results show that a nearly continuous sequence of *A. palmata* was recovered between 7,800 yr BP and 17,100 yr BP.

To help guide site selection three radiocarbon ages were determined while drilling was in progress. The remainder were

Barbados Offshore Drilling Program

Research Vessel N.U.S.C. Ranger

Cruise 88-13

18/November/88 - 6/December/88

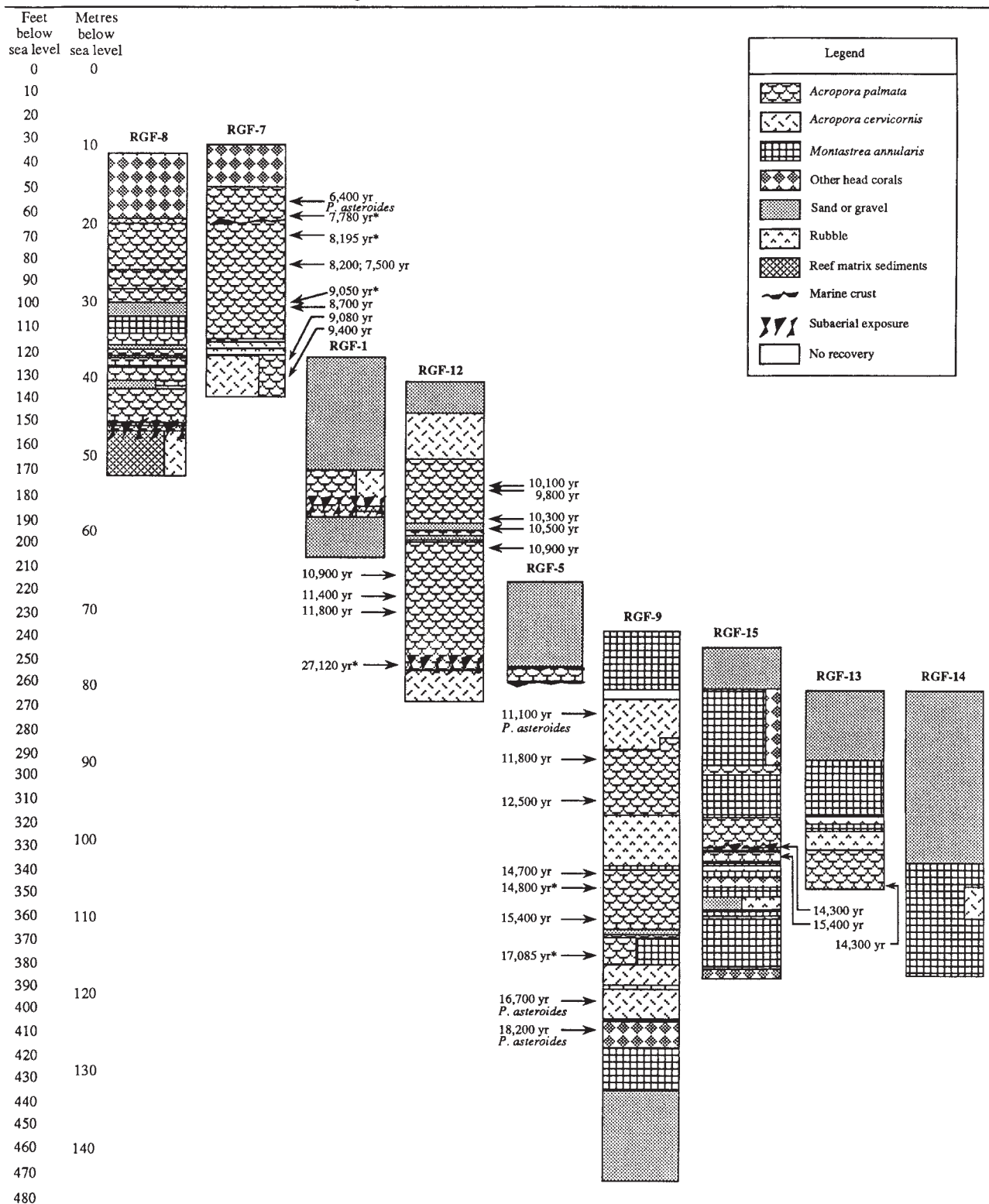


FIG. 1 Stratigraphic description of Barbados cores plotted against the modern sea level. Radiocarbon ages (yr) of individual corals are indicated by arrows. These ages are corrected for local seawater $\Delta^{14}\text{C}$ (the difference between the ^{14}C content of air and surface water) by subtracting 400 yr from the measured radiocarbon ages. All dated corals are *A. palmata* with four exceptions as noted where *P. asteroides* were dated. Sample at 25 m was dated twice and the 700-yr difference is assumed to arise from a ^{14}C

laboratory error. The younger age is assumed to be incorrect because it is stratigraphically inconsistent with neighbouring samples. This value is therefore not included in Fig. 2. Ages marked by an asterisk were measured at Geochron Laboratory, all others at Lamont-Doherty. Samples below the subaerial exposure horizons in cores RGF-8 and RGF-1 have been dated at >44,000 radiocarbon years.

determined after the cruise using conventional liquid scintillation counting or gas proportional counting. The average analytical uncertainty (1σ) of these measurements is ± 130 yr for ages less than 17,000 yr. All radiocarbon ages are corrected for the apparent age of sea water at Barbados by subtracting 400 yr (ref. 8). Radiocarbon ages are not corrected for secular changes in the atmospheric ^{14}C which may be large during the deglaciation⁹. As part of a companion study, the $^{230}\text{Th}/^{234}\text{U}$ ages of these samples are being determined by mass spectrometry¹⁰ to establish time-dependent changes in the surface ocean ^{14}C levels.

The deepest radiocarbon-dated sample of *A. palmata* occurs at 113.8 m below the present sea level (119.6 m if corrected for assumed mean uplift rate) and is dated to be 17,100 yr old. This age marks the end of the glacial period according to accelerator mass spectrometry (AMS) ^{14}C -dated $\delta^{18}\text{O}$ records from deep-sea cores^{11,12}. A sample of *Porites asteroides* from 124 m below the present sea level is 18,200 yr old. This date provides a maximum sea level estimate of 131 m (uplift corrected) for the last glacial maximum. Unlike *A. palmata*, *P. asteroides* is not restricted to shallow depths so this glacial maximum sea level estimate must be taken as an extreme. Extrapolation of the sea level curve to 18,000 yr BP indicates a sea level of 121 ± 5 m below the present level during the last glacial maximum, the best estimate so far.

Between 17,100 and 12,500 yr BP, the sea level increased by 20 m (Fig. 2). This first phase of deglaciation was terminated by an exceedingly rapid sea level rise of 24 m in less than 1,000 radiocarbon years, termed melt-water pulse IA (mwp-IA). Melt-water pulse IA (12,000 yr BP) corresponds to the latter part of Termination Ia (14,500–11,500 yr BP) as defined in deep-sea $\delta^{18}\text{O}$ records (refs 4, 11, later modified in ref. 12). Melt-water pulse IA and Termination Ia overlap to some extent, but cannot be synonymous. The rate of sea level rise was at a minimum at 11,000 yr BP, marking the beginning of the Younger Dryas chronozone, and remained low until 10,500 yr BP. The last half of the Younger Dryas chronozone (10,500–10,000 yr BP) is marked by increasing rates of sea level change, culminating in melt-water pulse IB (mwp-IB) centred at 9,500 yr BP. During mwp-IB, which closely corresponds to Termination Ib in deep-sea terminology, the sea level rose ~ 28 m, similar to the sea level rise associated with mwp-IA.

Comparison of seasonal insolation changes

According to the Milankovitch theory¹³ the Pleistocene ice ages were caused primarily by changes in the seasonal distribution of incoming radiation associated with orbit variations. The Northern Hemisphere ice sheets accumulate in the 40° – 70° N latitude range whereas the Antarctic ice sheet is located poleward of 70° S. Insolation changes associated with the 41,000-yr tilt cycle increase poleward in both hemispheres, whereas the 23,000-yr precessional cycle dominates insolation changes at middle and low latitudes and is out of phase between hemispheres. Since the work of Milankovitch, the time-dependent patterns of radiation have been calculated more precisely¹⁴. The Barbados sea level provides the first direct and high-resolution comparison between the astronomical changes and continental ice response. In Fig. 3, the changes in insolation over the Laurentide/Fennoscandinavian ice sheet and the Antarctic ice sheet are compared with the melt-water discharge rates calculated from the Barbados sea level curve. The discharge rates are plotted in radiocarbon years and must be converted to calendar years for comparison with the insolation curves¹⁵. Initial $^{230}\text{Th}/^{234}\text{U}$ dates by mass spectrometry indicate that the two melt-water pulses are not artefacts arising from distortions of the radiocarbon timescale (E. Bard, B. Hamelin and R.G.F., manuscript in preparation). The $^{230}\text{Th}/^{234}\text{U}$ age of mwp-IA is somewhat older, however, than the calibration curve¹⁵ would indicate. Although the radiocarbon-age calibration is not well established beyond 9,000 yr BP (ref. 15), the absolute-age estimates for the two melt-water pulses are 13,070 and 10,445 calendar years BP based on the ^{14}C calibration of ref. 15. This

comparison shows that the most intense interval of ice-sheet disintegration (13,070 calendar years BP) predates the Northern Hemisphere insolation maximum (centred at 11,000 years BP) by $\sim 2,000$ yr (Fig. 3). Melt-water pulse IB is 500 years younger than the insolation maximum. At present there is no astronomical explanation for separate melt-water pulses. The distribution of Northern Hemisphere ice and the retreat of its margins was not uniform¹⁶, and the relative contribution and timing of Antarctic ice decay is not known. These as well as other factors, such as changing ocean and atmosphere circulation patterns, may explain the variable discharge rates.

The global oxygen isotope curve

The ^{14}C -dated Barbados sea level curve can be scaled to $\delta^{18}\text{O}$ using the calibration of 0.011‰ per metre change in sea level⁵ (Fig. 2). This calibration is a maximum value and corresponds to a mean glacial-ice $\delta^{18}\text{O}$ value of -42% (standard mean ocean water). Based on ice-core $\delta^{18}\text{O}$ data from Greenland and Antarctica and modern $\delta^{18}\text{O}$ measurements in precipitation, this value provides a reasonable upper limit for mean ice composition for the entire deglaciation. The global $\delta^{18}\text{O}$ curve predicted from the Barbados sea level curve represents the ice volume component common to all $\delta^{18}\text{O}$ records from deep-sea cores. After taking into account the mixing time of the ocean, the global $\delta^{18}\text{O}$ curve can be subtracted from any high-resolution AMS ^{14}C -dated $\delta^{18}\text{O}$ curve to compare regional differences in the $\delta^{18}\text{O}$ of water and temperature.

Ocean $\delta^{18}\text{O}$ record of melt-water discharge rates

During mwp-IA and mwp-IB, glacial melt water discharged (primarily into the North Atlantic surface ocean) at maximum rates of $14,000 \text{ km}^3 \text{ yr}^{-1}$ and $9,500 \text{ km}^3 \text{ yr}^{-1}$ respectively (Fig. 3).

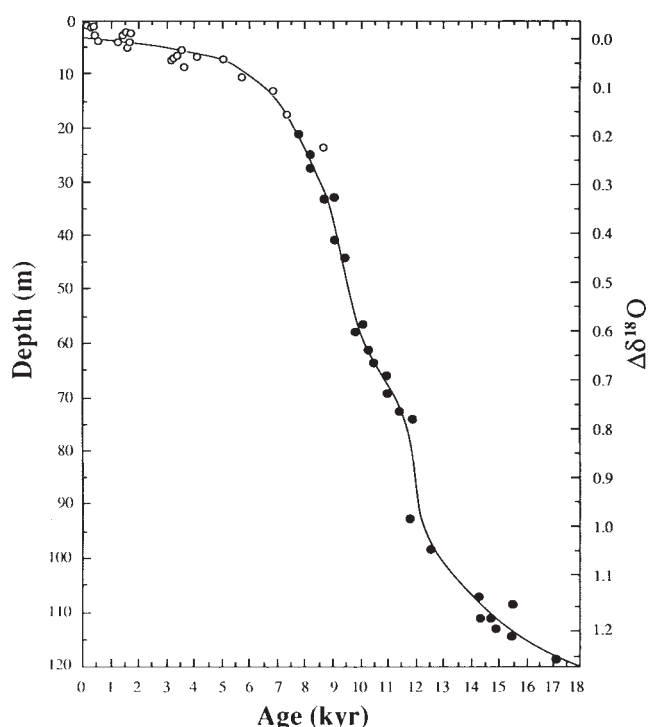


FIG. 2 Barbados sea level curve based on radiocarbon-dated *A. palmata* (filled circles) compared with *A. palmata* age-depth data¹ (open circles) for four other Caribbean island locations. All radiocarbon ages in this figure are corrected for local seawater $\Delta^{14}\text{C}$ by subtracting 400 yr from the measured radiocarbon ages⁸ but they are not corrected for secular changes in atmospheric ^{14}C levels¹⁵. The Barbados data are corrected for the estimated mean uplift of 34 cm kyr^{-1} . The right-hand axis of the Barbados sea level curve (solid line) is scaled to the estimated $\delta^{18}\text{O}$ change of mean ocean water using the calibration in ref. 5.

These glacial-melt-water rates are comparable to the modern fresh-water discharge rate ($11,400 \text{ km}^3 \text{ yr}^{-1}$) into the entire North Atlantic ($0-90^\circ \text{ N}$ latitude)¹⁷. During the first half of the Younger Dryas chronozone between 11,000 and 10,500 yr BP, melt-water discharge averaged $2,700 \text{ km}^3 \text{ yr}^{-1}$, a factor of five less than discharge rates during mwp-IA and at least a factor of three less than rates during mwp-IB. By comparison, the modern discharge rates for the Mississippi and St Lawrence Rivers are $560 \text{ km}^3 \text{ yr}^{-1}$ and $330 \text{ km}^3 \text{ yr}^{-1}$ respectively¹⁷.

The modern river discharge and marine precipitation is redistributed within the North Atlantic by ocean surface currents and results in strong oxygen isotope gradients of sea water in the north and north-west Atlantic bordering the subtropical gyre. Because glacial ice has very low $\delta^{18}\text{O}$ values relative to sea water, it is reasonable to assume that North Atlantic surface-water $\delta^{18}\text{O}$ distribution was dramatically affected by the sharp increase in seasonal melt-water discharge during mwp-IA and mwp-IB.

The oxygen isotopic composition of planktonic foraminifera is a function of both the ambient seawater oxygen isotopic composition and temperature. Foraminiferal species have different seasonal, geographic and vertical distributions whereas melt-water discharge is very seasonal. The most extensively dated (AMS ^{14}C) foraminiferal oxygen-isotope record in the North Atlantic is from piston core SU81-18 taken at a depth of 3,135 m near the coast of Portugal ($37^\circ 46' \text{ N}$, $10^\circ 11' \text{ W}$)¹². The oxygen isotope record of *G. bulloides* shows four distinct steps followed by $\delta^{18}\text{O}$ minima dated at 13,600, 12,300, 8,800 and 6,800 yr BP (Fig. 4). The two largest steps coincide with mwp-IA and mwp-IB. Although it is possible that these $\delta^{18}\text{O}$ shifts with minima at 12,300 and 8,800 yr BP are somehow related to local temperature changes, their coincidence with the salinity ($\delta^{18}\text{O}$) minima predicted by the Barbados sea level curve is strong evidence for regional salinity ($\delta^{18}\text{O}$) changes. I note that lower

salinity and warmer temperatures both result in reduced $\delta^{18}\text{O}$ values of foraminiferal tests. Assuming an average glacial melt-water $\delta^{18}\text{O}$ value of -42‰ , then the $\delta^{18}\text{O}$ of North Atlantic surface water will change at a corresponding rate of -1.2‰ for each 1% decrease in salinity. Therefore, the $\delta^{18}\text{O}$ minima in core SU81-18 (Fig. 4) can be produced by decreases in salinity of $<1\text{‰}$. The sea level increased only 10 m between 14,300 and 12,500 yr BP, an interval that spans the oldest $\delta^{18}\text{O}$ minima (13,600 yr) in core SU81-18.

Origin of the Younger Dryas event

In the preceding scenario, the Younger Dryas chronozone in the marine realm is characterized by minimal glacial-melt-water discharge, particularly between 11,000 and 10,500 yr BP, and is distinguished as an event primarily by its temporal position between two episodes of lower surface salinity. This interpretation of the origin of the Younger Dryas event is fundamentally different from the explanations involving the polar-front movements linked to ice-sheet elevation^{2,3,12,18-20}. It is the exact opposite of explanations for the Younger Dryas that call upon maximum influxes of glacial ice and water between 11,000 and 10,000 yr BP¹⁸⁻²².

Apparently there are very different regional responses to the melt-water pulses. For example, Mangerud *et al.*²³ have shown that in western Norway there were two major glacial re-advances into the North Sea at 12,200 and 10,000 yr BP. In the north-east Atlantic, unusually high abundances of the polar species *N. pachyderma sin* are found in association with the Vedde ash layer dated at 10,600 years BP (ref. 23). This peak abundance of polar species has been interpreted as a Younger Dryas cold event in the north-east Atlantic¹⁸⁻²⁰ although it apparently coincides with a low melt-water discharge rate. Ruddiman *et al.*⁸ concluded that the Vedde ash, such as found at core site V23-81 on the Feni Drift, was delivered to the north-east Atlantic

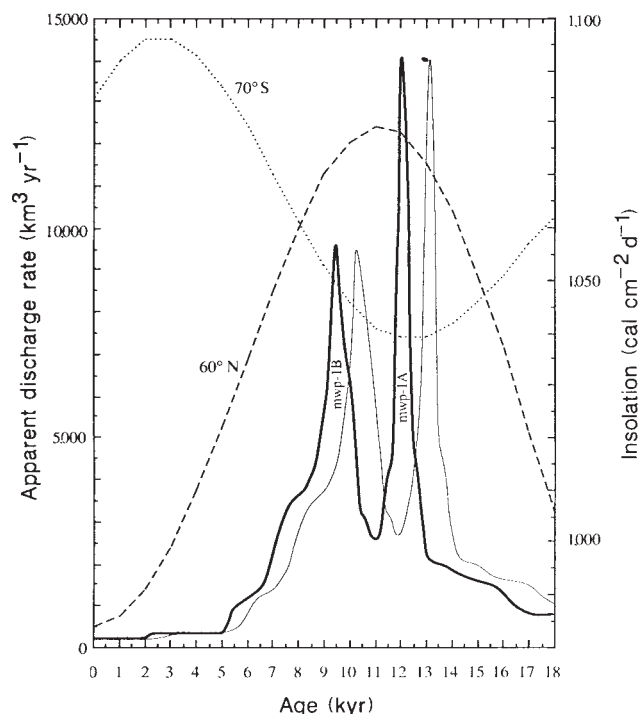


FIG. 3 The rate of glacial melt-water discharge calculated from the Barbados sea level curve compared to summer insolation. The discharge curve (thick line) is plotted according to radiocarbon years uncorrected for secular changes in atmospheric ^{14}C . Converting radiocarbon years BP to calendar years BP using the calibration in ref. 15 shifts the ages of mwp-IB and mwp-IA as shown by the thin line. Our preliminary $^{230}\text{Th}/^{234}\text{U}$ dates indicate that this correction is too small for mwp-IA (E. Bard, B. Hamelin and R.G.F., manuscript in preparation). By comparison, the insolation at 60° N (dashed line) and 70° S latitude (dotted line) are plotted against absolute time¹⁴.

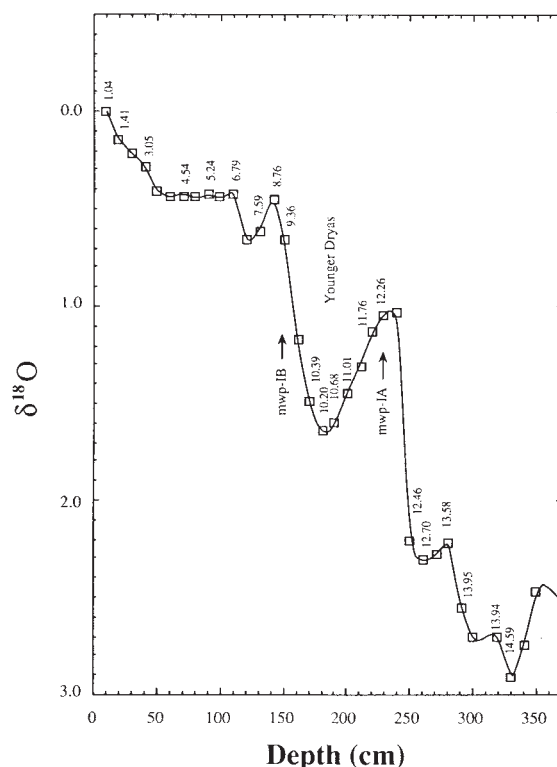
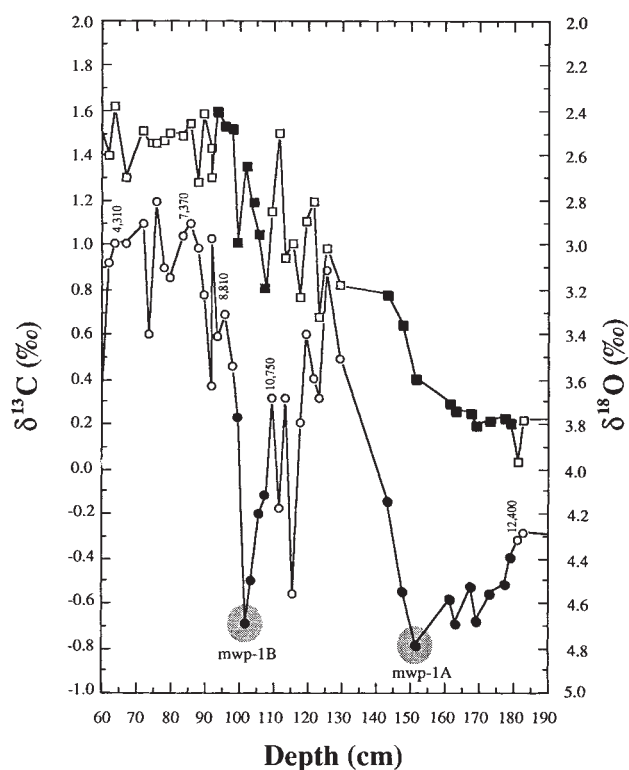


FIG. 4 The AMS ^{14}C -dated oxygen isotope record of *G. bulloides* from core SU81-18 located offshore of Portugal^{12,34}. The AMS- ^{14}C ages have been corrected for local seawater $\Delta^{14}\text{C}$ by subtracting 400 yr from the measured radiocarbon ages⁸. The estimated depth of mwp-IA and mwp-IB are marked by arrows.

FIG. 5 Oxygen and carbon isotope data from core EN120GGC1²⁹ for the last deglacial interval. Age assignments are based on four AMS-¹⁴C dates from a companion core that has been stratigraphically correlated to EN120GGC1³⁰. Using this age model, $\delta^{13}\text{C}$ minima are approximately coincident with mwp-1A and mwp-1B. Closed symbols mark intervals around mwp-1A and mwp-1B and illustrate that $\delta^{13}\text{C}$ minima occur during the unidirectional shifts in $\delta^{18}\text{O}$.



by sea ice originating in the Norwegian Sea. Recent sea-ice studies²⁴ show, however, that it is plausible that the *N. pachyderma sin* was also imported to the north-east Atlantic by entrapment in sea ice that had formed in the Norwegian Sea.

Lower surface-water salinities during mwp-1A and 1B are not inconsistent with the faunal distributions in the north-east Atlantic described in, for example, ref. 18. Polar station PAPA, at 50° N latitude in the northeastern Pacific, has a low-salinity mixed layer and may be a good analogue for the North Atlantic at 12,000 and 9,500 yr BP. Today at station PAPA, an 8 °C seasonal warming of the thin mixed layer results in a seasonal succession of planktonic foraminiferal species from polar to subpolar assemblages²⁵. Lower salinity might also help explain the evidence for warmer climates on the southern European continent during the Bølling-Allerød and Preboreal periods (see ref. 12 for a review). Lower surface salinities generally result in a thinner surface mixed layer with less thermal inertia. Energy-balance climate models show²⁶ that a thinner surface mixed layer in the North Atlantic (>40° N) leads to warmer surface temperatures over land in summer and fall. Much larger regional responses over Western Europe are predicted in three-dimensional global-climate-model simulations²⁶. Conversely, the lack of melt-water stratification of North Atlantic surface waters during the Younger Dryas would allow a return to cooler deglacial conditions recorded by planktonic foraminifera in the North Atlantic ~10,600 yr BP (refs 19, 20, 27) and by western European continental vegetation.

Broecker *et al.*^{27,28} dated the minimum $\delta^{18}\text{O}$ value (maximum melt-water discharge) in the Gulf of Mexico cores EN32-PC6 and EN32-PC4 at 11,840 and 12,000 yr BP, which are equivalent in age to the AMS ¹⁴C-dated $\delta^{18}\text{O}$ minimum in North Atlantic core SU81-18 (12,260 yr BP)¹². These two events correspond precisely with mwp-1A, dated at 12,000 yr BP (Fig. 3). Broecker *et al.*^{27,28} noted that the maximum glacial-melt-water discharge from the Mississippi River was followed by minimal discharge between 10,500 (EN32-PC6) and 11,100 yr BP (EN32-PC4). They reasoned that this reduction in Mississippi discharge reflected a shift in the drainage from the Mississippi to the St Lawrence

River and was not related to changes in glacial melting rates. They further speculated that the melt-water diversion caused a salinity decrease in the North Atlantic strong enough to result in a decrease in North Atlantic deep-water formation during the Younger Dryas. A reduction in the heat released from the ocean as the northward advection of subtropical surface waters ceased was associated with the reduced formation of North Atlantic deep water. The discharge curve derived from the Barbados sea level curve predicts a minimum in Mississippi discharge 11,000 yr BP and therefore it is unnecessary to call upon the unusual combination of events required by the model in refs 27, 28. Furthermore, the discharge curve may explain why no melt-water spike was ever identified in the North Atlantic oxygen isotope records corresponding to an assumed diversion of melt water during the Younger Dryas event.

Deep-water response to melt-water pulse

Broecker *et al.*^{27,28} cite Cd/Ca and $\delta^{13}\text{C}$ data reported from Bermuda Rise core EN120GGC1²⁹ as evidence that melt water shut down the production of North Atlantic deep water during the Younger Dryas. Boyle and Keigwin's²⁹ age model for the deglacial section of core EN120GGC1 is based on linear interpolation between age assignments of 14,000 yr BP for Termination Ia and 8,500 yr BP for Termination Ib as well as correlation of the %CaCO₃ record with that of a neighbouring core that has two AMS ¹⁴C dates on the deglaciation. These ages for Terminations Ia and Ib were taken from the AMS ¹⁴C dates on North Atlantic core CH73-139C¹¹. AMS ¹⁴C studies on other cores have shown that core CH73-139C suffers from considerable bioturbation¹². The revised ages for Termination Ia are 14,500–12,260 yr BP and for Termination Ib, 10,390–9,360 yr BP (ref. 12).

Applying a more accurate chronology to core EN120GGC1³⁰ shows that the two intervals of highest melt-water discharge are also characterized by extremely low $\delta^{13}\text{C}$ values, presumably because of reduced influence of North Atlantic deep water at this site (Fig. 5). The section corresponding to Younger Dryas interval begins with high $\delta^{13}\text{C}$ values which then decrease to a minimum value during Termination 1b. If this revised timescale

for core EN120GGC1 is accurate, then the Younger Dryas and Preboreal chronozones are so condensed that bioturbation must be considered as a significant influence. The fact that a relatively subtle adjustment of age assignments can alter the interpretation of deep ocean response to deglaciation emphasizes the necessity for AMS ^{14}C dating of splits of the same samples used for stable isotope and Cd/Ca measurements in core EN120GGC1. Until this is done, it is equally plausible to suggest that the production of North Atlantic deep water was reduced or shut down within the Bølling/Older Dryas (mwp-IA) and Preboreal (mwp-IB) chronozones and not during the Younger Dryas.

Continental record of melt-water discharge rates

The rapid disintegration of the Northern Hemisphere ice sheets during mwp-IA and mwp-IB must have introduced large volumes of low- $\delta^{18}\text{O}$ melt water into the North Atlantic during the summer season. This $\delta^{18}\text{O}$ -spiked melt water has two possible fates: entrainment in North Atlantic deep water or evaporation¹⁷. Because deep-water formation may have been halted briefly during the peak discharge, evaporation of $\delta^{18}\text{O}$ -spiked surface water must have occurred to some extent.

The Greenland ice cores Dye 3³¹⁻³³ and Camp Century are both marked by two distinct $\delta^{18}\text{O}$ minima during deglaciation (Fig. 6). Many researchers assumed that these $\delta^{18}\text{O}$ minima indicated cooler temperatures²⁹⁻³² and that they were associated with a southward migration of the North Atlantic polar front^{19,20}. However, recent ^{14}C dating of air bubbles in the Dye 3 ice core³³ enables a direct comparison to the ^{14}C dated Barbados sea level curve (Fig. 6). Six AMS ^{14}C measurements were made on samples taken at depths between 1,660 and 1,780 m (ref. 33). Using the age-depth relationship (squared correlation coefficient = 0.93) determined from these data, the estimated age for the $\delta^{18}\text{O}$ minimum at 1,792 m (Fig. 6) is 9,320 ^{14}C -years BP. The age of this $\delta^{18}\text{O}$ minimum in Dye 3 matches the age for mwp-IB (9,500 ^{14}C -years BP). It is not clear whether the ^{14}C age is concordant with the estimated calendar age based upon annual layer counts³¹, because the radiocarbon timescale is not well calibrated in this time range. I believe that these presumed temperature minima in the Greenland $\delta^{18}\text{O}$ records may in fact be the result of source-water variability as well as temperature. If so, they may be the terrestrial equivalents of $\delta^{18}\text{O}$ minima in deep-sea cores such as SU81-18 (Fig. 4).

The immediate implications of the Barbados sea level record compel us to re-evaluate some of the traditional depictions of

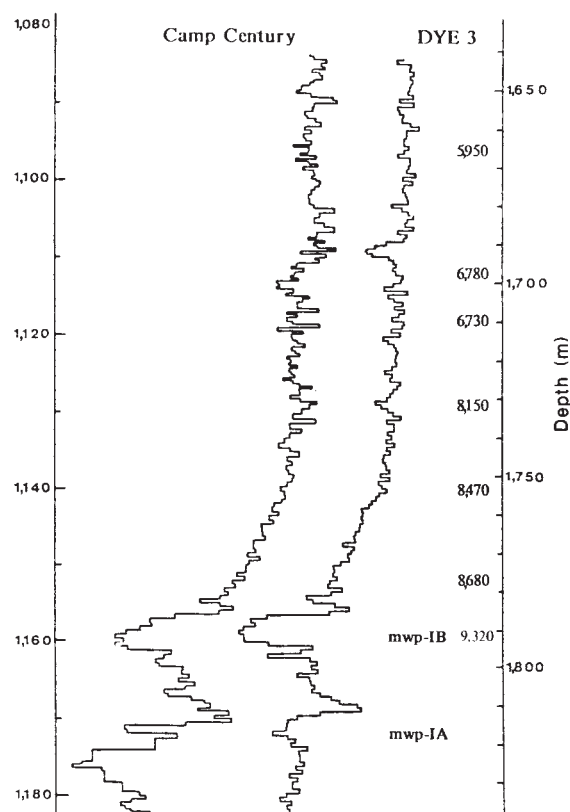


FIG. 6 Plot of terrestrial $\delta^{18}\text{O}$ records from Greenland ice cores Dye 3 and Camp Century^{31,32} with six radiocarbon ages³³ indicated for Dye 3. Age for mwp-IB is estimated from the ^{14}C -depth relationship³³.

the Younger Dryas event. Whether or not the reinterpretations presented here are correct depends on the reliability of correlations between radiocarbon-dated marine and terrestrial signals. My results emphasize the primary role of melt-water discharge rates on the production of North Atlantic deep water and rapid climate change in western Europe. □

Received 18 May; accepted 19 October 1989.

- Lighty, R. G., MacIntyre, I. G. & Stuckenrath, R. *Coral Reefs* **1**, 125-130 (1982).
- Ruddiman, W. F. in *North American and Adjacent Oceans During the Last Deglaciation* (eds Ruddiman, W. F. & Wright, H. E. Jr) **K-3** 463-478 (Geol. Soc. Am., Boulder, 1987).
- Ruddiman, W. F. in *North American and Adjacent Oceans During the Last Deglaciation* (eds Ruddiman, W. F. & Wright, H. E. Jr) **K-3** 137-154 (Geol. Soc. Am., Boulder, 1987).
- Duplessy, J. C., Delibrias, G., Turon, J. L., Pujol, C. & Duprat, J. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 121-144 (1981).
- Fairbanks, R. G. & Matthews, R. K. *Quat. Res.* **10**, 181-196 (1978).
- MacIntyre, I. G. *Am. Ass. Petrol. Geol. Bull.* **56**, 720-738 (1972).
- Clark, J. A., Farrell, W. E. & Peltier, W. R. *Quat. Res.* **9**, 265-287 (1978).
- Bard, E. *Palaeoceanography* **3**, 635-645 (1988).
- Oeschger, H. et al. *Radiocarbon* **22**, 299-310 (1980).
- Edwards, R. L., Chen, J. H., Ku T.-L. & Wasserburg, G. L. *Science* **236**, 1547-1553 (1987).
- Duplessy, J.-C. et al. *Nature* **320**, 350-352 (1986).
- Bard, E. et al. *Nature* **328**, 791-794 (1987).
- Milankovitch, M. M. *Canon of Insolation and The Ice-Age problem* (Königlich Serbische Akademie, Beograd, 1941). English translation published by the U.S. Department of Commerce and the NSF, Washington, DC.
- Berger, A. L. *Quat. Res.* **35**, 215-240 (1978).
- Stuiver, M. et al. *Radiocarbon* **28**, 969-979 (1986).
- Andrews, J. T. in *North American and Adjacent Oceans During the Last Deglaciation* (eds Ruddiman, W. F. & Wright, H. E. Jr) **K-3** 13-38 (Geol. Soc. Am., Boulder, 1987).
- Baumgartner, A. & Reichel, E. *The World Water Balance* (Elsevier, New York, 1975).
- Ruddiman, W. F., Sancetta, C. D. & McIntyre, A. *Phil. Trans. R. Soc. B* **280**, 119-142 (1977).

- Ruddiman, W. F. & McIntyre, A. *Science* **212**, 617-627 (1981).
- Ruddiman, W. F. & McIntyre, A. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **35**, 145-214 (1981).
- Mercer, J. H. *Arctic Alpine Res.* **1**, 227-234 (1969).
- Johnson, R. G. & McClure, B. T. *Quat. Res.* **6**, 325-353 (1976).
- Mangerud, J., Lie, S. E., Furnes, H., Kristiansen, I. L. & Lomo, L. *Quat. Res.* **21**, 85-104 (1984).
- Spindler, M. & Diekmann, G. S. *Polar Biol.* **5**, 185-191 (1986).
- Reynolds, L. & Thunell, R. C. *J. Foraminiferal Res.* **15**, 282-301 (1985).
- Schneider, S. H., Peteet, D. M. & North, G. R. in *Abrupt Climate Change* (eds Berger, W. H. & Labeyrie, L. D.) 399-418 (Reidel, Dordrecht, 1987).
- Broecker, W. S. et al. *Palaeoceanography* **3**, 1-19 (1988).
- Broecker, W. S. et al. *Nature* **341**, 318-321 (1989).
- Boyle, E. A. & Keigwin, L. *Nature* **330**, 35-40 (1987).
- Keigwin, L. D., Jones, G. A. *Deep Sea Res.* **36**, 845-867 (1989).
- Hammer, C. U., Clausen, H. B. & Tauber, H. *Radiocarbon* **28**, 284-291 (1986).
- Dansgaard, W., White, J. W. & Johnson, S. J. *Nature* **339**, 532-534 (1989).
- Andree, M. et al. *Radiocarbon* **28**, 417-423 (1986).
- Bard, E. *Quat. Res.* **31**, 381-391 (1989).

ACKNOWLEDGEMENTS. For help with various aspects of the cruise and technical assistance I thank W. Svendsen (Longyear Co.), P. Gemeinhardt, K. Taylor and his drilling crew (Fugro McClelland Inc.), Captain M. Bordeaux, Firstmate P. Scott and the crew of the RV *Ranger*, R. Miller, T. Baker, C. Charles, T. Janacek, R. Lighty, D. Oppo, C. Ravelo, L. Barker, W. Hunte, P. McConney and Mr. Hope. I also thank G. Mathieu for the ^{14}C age determinations, E. Bard, W. Broecker, G. Kukla, E. Lighty, R. Peltier, W. Ruddiman and J. Wright for discussions, L. Keigwin and J. C. Duplessy for reviews and R. Matthews for his continued cooperation and advice on the Pleistocene history of Barbados. This research was supported by the NSF.

Cloning by functional expression of a member of the glutamate receptor family

Michael Hollmann, Anne O'Shea-Greenfield, Scott W. Rogers & Stephen Heinemann

Molecular Neurobiology Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

We have isolated a complementary DNA clone by screening a rat brain cDNA library for expression of kainate-gated ion channels in *Xenopus* oocytes. The cDNA encodes a single protein of relative molecular mass (M_r) 99,800 which on expression in oocytes forms a functional ion channel possessing the electrophysiological and pharmacological properties of the kainate subtype of the glutamate receptor family in the mammalian central nervous system.

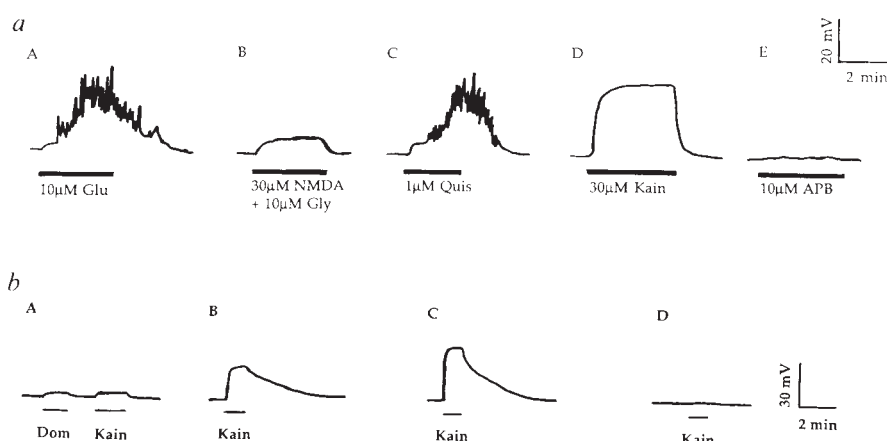
THE synapse plays a key role in the nervous system, and it is likely that biochemical changes at the synapse underlie some aspects of higher brain function. Most plausible theories of learning, for example, depend upon changes in the efficiency of chemical synapses^{1,2}. Glutamate receptors are believed to play a central part both in memory acquisition and the establishment of "topographical maps"^{3,4}. Hence, this receptor family has become the focus of much experimental work. Glutamate receptors are found throughout the mammalian brain and represent the predominant excitatory neurotransmitter system⁵. The best-studied glutamate receptors are ligand-gated ion channels that are permeable to cations. Based on pharmacological and electrophysiological data these receptors have been classified into four main subtypes: NMDA (*N*-methyl-D-aspartate), quisqualate, kainate and 2-amino-4-phosphonobutyrate (APB).

Recently, a G protein-coupled glutamate receptor representing a fifth class with a very different mechanism of action has been identified⁶. The fact that the subtypes show differences in their distribution in the brain⁷ indicates that they represent distinct gene products with unique structures and functions.

Despite the prominent part these receptors play in normal synaptic transmission as well as in neuronal plasticity, their molecular characteristics have remained elusive, mainly because of a lack of ligands that bind irreversibly and with high specificity. Conventional biochemical approaches to the isolation of these receptors have so far either not succeeded in progressing beyond crude receptor solubilization, as is the case for the NMDA⁸ and quisqualate⁹ subtypes, or have resulted in the isolation of ligand-binding proteins for kainate¹⁰⁻¹² or glutamate^{13,14} whose functions remain unclear. To circumvent these problems, we adopted a different approach involving the use of an oocyte expression system¹⁵ to guide the selective cloning of cDNAs encoding functional glutamate-gated ion channels.

Here, we report the isolation of a cDNA clone, GluR-K1, from a rat forebrain cDNA library. This clone, on transcription *in vitro* and injection into *Xenopus* oocytes, expresses a functional glutamate receptor with the pharmacological properties of the kainate subtype. The primary structure of the protein encoded by the GluR-K1 cDNA has been deduced, and it is compared with the sequences of other ligand-gated ion channels, the nicotinic acetylcholine receptors (nAChR), and the glycine and γ -aminobutyric acid (GABA_A) receptors. The pattern of distribution of GluR-K1 mRNA is described and compared with kainate binding sites in the brain.

FIG. 1 *a*, Voltage traces obtained from a single *Xenopus* oocyte injected with 75 ng poly(A)⁺ RNA isolated from rat forebrain 2 days before recording. Bars indicate duration of superfusion with indicated agonists. Glu, L-glutamate; NMDA, *N*-methyl-D-aspartate; Gly, glycine; Quis, quisqualate; Kain, kainate; APB, 2-amino-4-phosphonobutyrate. *b*, Voltage recordings from *Xenopus* oocytes 10 days (trace A) or 3 days (traces B–D) after injection with 25 ng λ ZAP RTB1 sense *in vitro* RNA (total library of 850,000 independent clones) (A), 25 ng λ ZAP-GluR-K1 sense *in vitro* RNA (B), 1.25 ng of pGluR-K1 sense *in vitro* RNA (C) or 1.25 ng of pGluR-K1 antisense *in vitro* RNA (D). Dom, domoate (10 μ M); Kain, kainate (100 μ M). Bars indicate duration of superfusion with agonist. **METHODS.** RNA was isolated from rat total forebrain by the guanidine isothiocyanate method³⁹ and subsequently poly(A)⁺ selected⁴⁰. The Uni-ZapTM XR cloning kit from Stratagene was used to construct a directional cDNA library of 850,000 independent clones in the bacteriophage expression vector λ ZAP II, where the insert clones are flanked by T3 and T7 RNA polymerase promoters at their 5' and 3' ends, respectively, allowing sense as well as antisense transcription⁴¹. DNA from pools of phage clones was linearized with Xho I, and diguanosine triphosphate-capped RNA was synthesized *in vitro* using the RNA transcription kit from Stratagene. A DNase I digest was performed to remove the template after synthesis. *Xenopus* oocytes were prepared as previously described³³ and injected with 1–



100 ng RNA, usually in a volume of 50 nl. Injected oocytes (20 per batch of RNA to be tested) were maintained in daily changed Barth's saline³³ at 17 °C. During the initial stages of the library screening, oocytes were checked for responses after 3, 5, 7, 9 and 10 days. During later stages of the screening, recordings were made on days 3 and 5 after injection. To facilitate rapid, sensitive screening of large numbers of oocytes we used voltage recording rather than voltage clamp, as described previously³³. Depolarizations as small as 1–2 mV could be measured reliably.

Isolation of clone λ ZAPII-GluR-K1

Xenopus oocytes were injected with poly(A)⁺ RNA isolated from rat forebrain, 2–10 days after which the oocytes were tested electrophysiologically for their ability to respond to selective agonists for glutamate receptor subtypes. Depolarizations induced by both glutamate and quisqualate displayed a biphasic pattern composed of a fast-acting, smooth, presumably ligand-gated, ion channel response and a longer lasting, fluctuating response that was probably second-messenger mediated¹⁶ (Fig. 1a, traces A, C). NMDA and kainate elicited smooth responses with fast onsets (Fig. 1a, traces B, D). APB gave no response (Fig. 1a, trace E).

A directional cDNA library (λ ZAPII RTB1; complexity, 8×10^5 elements) consisting of 18 independent sublibraries each of 44,000 clones was constructed from this poly(A)⁺ RNA using the bacteriophage expression vector λ ZAPII. A pool of *in vitro* transcripts made separately from all 18 amplified sublibraries was injected into oocytes. Small depolarizations (1–3 mV) were seen in voltage recordings from oocytes challenged with 100 μ M kainate 10 days after injection (Fig. 1b, trace A). No responses to NMDA or quisqualate were detected. Neither normal oocytes nor water-injected oocytes showed any response to glutamate receptor agonists. Subsequently, pools of 44,000, 4,000, 400 and 40 clones were tested and in each of these tests at least one pool responded to kainate. To ensure throughout the screening procedure that the very small responses observed initially were not recording artefacts, we showed that: (1), responses in a given oocyte were reproducible; (2) responses were fast (within one second of agonist application); (3) responses were readily reversible on superfusion of the oocyte with control Ringer solution; and (4) domoate (10 μ M) gave a response similar to the one elicited by 100 μ M kainate. The pools yielding the largest responses at each stage were selected for further subdivision. The clones in the final positive pool of 40 clones were analysed

for their insert size, and the 12 clones with the largest inserts (all >2 kilobases (kb)) were tested individually for their ability to direct the synthesis of a functional kainate receptor. Only one clone, λ ZAPII-GluR-K1, carrying a 3-kb insert, was found to elicit kainate responses in oocytes (Fig. 1b, trace B).

Characterization

The plasmid pGluR-K1 was rescued from bacteriophage λ ZAPII-GluR-K1 and, on transcription *in vitro*, sense RNA (but not antisense RNA) induced kainate responses when injected into oocytes (Fig. 1b, traces C, D). As little as 10 pg of GluR-K1 sense transcript under voltage-clamp conditions (−70 mV holding potential) gave detectable responses to 100 μ M kainate. To rule out the possibility that GluR-K1 encoded a transcription factor, oocytes injected with pGluR-K1 *in vitro* transcripts were kept in medium supplemented with 50 μ g ml^{−1} actinomycin D to inhibit endogenous transcription¹⁷. These oocytes exhibited the same responses to kainate as those kept in control medium. Therefore, injection-induced transcription from the oocyte genome seems unlikely to contribute to the observed responses.

L-Glutamate evoked much smaller responses than did kainate, and even at 1 mM, elicited only 50% of the depolarization seen with 30 μ M kainate. This is consistent with previous reports that glutamate is only a weak agonist for the kainate receptor subtype³. Other glutamate receptor agonists such as NMDA, quisqualate and L-aspartate evoked no responses when applied at 150 μ M, 10 μ M and 100 μ M, respectively. Unrelated neurotransmitter receptor agonists such as glycine, GABA, serotonin and nicotine also failed to evoke responses, even when tested at concentrations as high as 1 mM. Glycine did not potentiate the kainate response (data not shown). Dose-response curves for kainate and domoate were recorded (Fig. 2b), and EC₅₀ (effector concentration for half-maximal response) values of

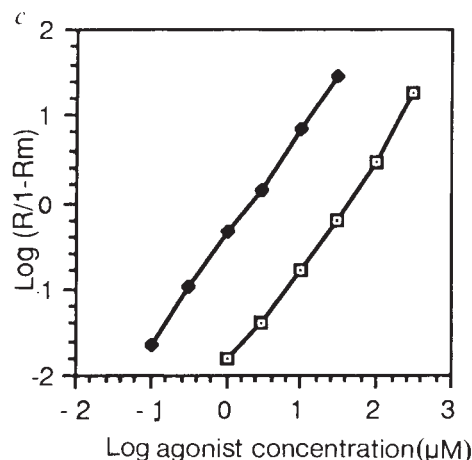
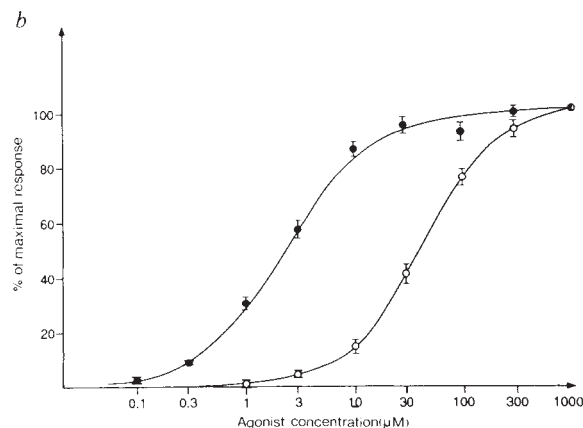
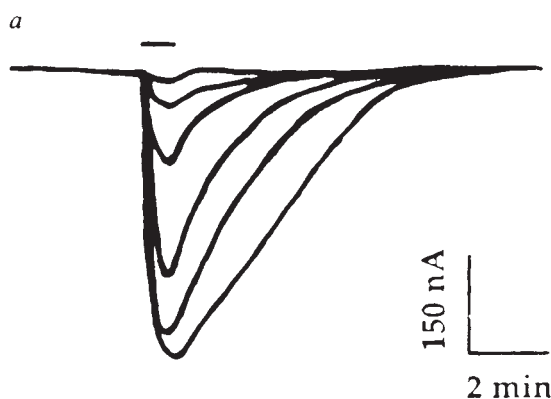


FIG. 2 *a*, Superimposed voltage clamp traces, recorded at a holding potential of −70 mV from a single oocyte injected 2 days before recording. The oocyte was injected with 1.25 ng pGluR-K1 *in vitro* synthesized RNA transcript, and challenged with increasing kainate concentrations of 1–1,000 μ M. The agonist was applied for 40 s, followed by control Ringer superfusion. Intervals between applications were 5 min. Note that the responses do not desensitize. *b*, Dose-response curves for kainic acid and domoic acid, recorded as in *a*; each point represents the mean $\pm$ s.e.m. of three responses, measured in different oocytes. The EC₅₀ values derived from these curves are 39 μ M and 1.8 μ M for kainic acid and domoic acid, respectively. Closed circles, domoate; open circles, kainate. *c*, Hill plots for oocyte responses to domoate (closed circles) and kainate (open circles), constructed using the mean values of the data from the dose-response curves of *b*. R_m, maximal response; R, response. Hill coefficients of 1.23 and 1.24 for kainate and domoate, respectively, were obtained.

METHODS. The insert cDNA of the bacteriophage clone λ ZAPII-GluR-K1 was cut out with *Eco*RI-*Xho*I, blunt-ended and subcloned into the *Sma*I site of the bacteriophage vector M13mp19 (ref. 42), yielding clones with both orientations of the cDNA. Overlapping deleted subclones were constructed for each strand orientation using the CycloneTM kit from USB. Single-strand sequencing by the dideoxynucleotide chain-termination method⁴³ was done with all (45) subclones, and 10 oligonucleotide primers were synthesized to facilitate sequencing across areas where compressions or gaps were encountered in the sequences derived from the deletion subclones. Complete sequences for both strands were thus obtained. IntelliGenetics software packages (IntelliGenetics version 5.0, and PC/GeneTM) were used for analysing sequences.

METHODS. The insert cDNA of the bacteriophage clone λ ZAPII-GluR-K1 was cut out with *Eco*RI-*Xho*I, blunt-ended and subcloned into the *Sma*I site of the bacteriophage vector M13mp19 (ref. 42), yielding clones with both orientations of the cDNA. Overlapping deleted subclones were constructed for each strand orientation using the CycloneTM kit from USB. Single-strand sequencing by the dideoxynucleotide chain-termination method⁴³ was done with all (45) subclones, and 10 oligonucleotide primers were synthesized to facilitate sequencing across areas where compressions or gaps were encountered in the sequences derived from the deletion subclones. Complete sequences for both strands were thus obtained. IntelliGenetics software packages (IntelliGenetics version 5.0, and PC/GeneTM) were used for analysing sequences.

39 μM and 1.8 μM , respectively, were derived. These values were in good agreement with results found in oocytes injected with total poly(A)⁺ RNA^{16,18}. Hill coefficients calculated from these curves (Fig. 2c) were 1.23 and 1.24 for kainate and domoate, respectively, indicating the possible cooperative action of several subunits. The average reversal potential extrapolated from current responses to 10 μM kainate obtained at a series of holding potentials between 0 and -130 mV was 10 mV. Responses to kainate did not desensitize, not even after prolonged (up to 10 min) superfusion with high concentrations (100 μM) of the agonist. These findings are also consistent with earlier reports from experiments with kainate receptors expressed from total poly(A)⁺ RNA¹⁹. Similarly, the pharmacological profile of GluR-K1 (see Table 1), as revealed by the inhibiting properties of various known glutamate receptor antagonists, is consistent with previous reports for kainate receptors in systems where total poly(A)⁺ RNA was used as a source of kainate receptor message^{16,18}. Of all the compounds tested (each at 1 mM), the glutamate receptor antagonist of broad specificity, kynurenic acid, was the most potent inhibitor of the kainate-evoked depolarizations in oocytes injected with GluR-K1 RNA synthesized *in vitro*. To a lesser extent, γ -D-glutamylglycine (γ -DGG), reported to block preferentially kainate and NMDA receptors but not quisqualate receptors²⁰, inhibited kainate responses, as did γ -D-glutamylaminomethylsulphonate (GAMS), which acts preferentially on kainate and quisqualate receptors²¹. Similarly, 2,3-*cis*-piperidine dicarboxylic acid (PDA), which is known to block all subtypes of glutamate receptors⁵, blocked the GluR-K1 response. Glutamate diethyl-ester (GDEE), which is thought to inhibit preferentially quisqualate-type glutamate receptors⁵, did not block the response significantly, but instead displayed weak agonist properties. The NMDA receptor antagonists AP5 (D-(-)-2-amino-5-phosphonovaleric acid) and CPP (3-(2-carboxypiperazin-4-yl)-propyl-1-phosphate), as well as NMDA itself, slightly inhibited the kainate responses, thus acting as weak antagonists without showing any agonist properties.

Taken together, the electrophysiological properties observed in oocytes injected with GluR-K1 transcript, as well as the observed pharmacological properties, strongly indicate that GluR-K1 represents a functional kainate receptor indistinguishable from the one observed in oocytes injected with total poly(A)⁺ RNA. Thus, it appears that the single protein subunit encoded by GluR-K1 is sufficient to form an active receptor-ion-channel complex. Whether this complex is composed of a single protein subunit *in vivo*, or whether it is a homo-oligomer assembled from several identical subunits, cannot be decided. Given the fact that the depolarization induced by RNA *in vitro* transcripts of the isolated clone GluR-K1 is smaller than the one observed with forebrain poly(A)⁺ RNA (Fig. 1), it seems

TABLE 1 Pharmacology of the glutamate receptor encoded by GluR-K1

Compound tested	Compound alone (%) [*]	Compound plus kainate (%) [†]
Kainate (agonist control)	100	100
Kynurenic acid	-3.4 ± 0.2	9.6 ± 1.3
γ -DGG	-0.1 ± 0.5	30.8 ± 1.1
GAMS	1.0 ± 0.6	30.7 ± 1.1
GDEE	24.8 ± 5.7	97.8 ± 6.6
PDA	2.5 ± 0.7	31.3 ± 2.0
APV	3.1 ± 1.4	73.7 ± 3.2
CPP	9.1 ± 0.7	78.8 ± 0.9

Oocytes had been injected with 1.25 ng GluR-K1 *in vitro* sense RNA 3 days before the recording. The oocytes were voltage-clamped at -70 mV, and the test compounds (all at 1 mM, except kainate, which was 30 μM) applied by rapid superfusion, with 5-min intervals between drugs. Peak currents were recorded, and each number represents the average of 3 recordings from 3 different oocytes, $\pm$ s.e.m. The 100% current response corresponds to 40–200 nA, depending on the oocyte.

^{*} Per cent of the response evoked by 30 μM kainate immediately before drug application.

[†] Per cent of the response seen with 30 μM kainate alone immediately before application of drug-kainate mixture.

possible that additional receptor subunits exist that might modify the channel properties. The slight deviation of the Hill coefficient from unity suggests the presence of interacting multiple subunits.

DNA and deduced amino-acid sequences

The cDNA insert of the plasmid pGluR-K1 was subcloned into M13mp19 and sequenced (Fig. 3). An open reading frame of 2,721 base pairs (bp) was found within the total length of 2,992 bp. Thus, the predicted protein consists of 889 amino acids, with a calculated M_r of 99,769. The deduced protein sequence contains a putative signal peptide of 18 amino acids at its N terminus, with a predicted cleavage site that conforms to empirical rules²². The N terminus of the protein therefore is expected to be located extracellularly. There are six possible N-glycosylation sites present in the presumed extracellular domain. Sequence comparisons with other ligand-gated ion channels, the nicotinic acetylcholine receptors, GABA_A receptors and the glycine receptor, reveal little overall homology and much less than anticipated in the light of the recently proposed 'superfamily' hypothesis^{23,24} of ligand-gated ion channels. In fact, using the sequence comparison algorithm of Feng, Johnson and Doolittle²⁵, no significant over-all homology is detected between the protein encoded by GluR-K1 and the neuronal neuronal nAChR subunits^{26,27} α 2, α 3, α 4, β 2, β 3 or β 4, GABA_A receptor subunits²⁸ α or β , or the glycine receptor subunit of M_r 48K²⁹.

All ligand-gated ion-channel subunits sequenced previously have a conserved extracellular region characterized by two cysteine residues spaced 14 amino acids apart, with conserved proline and aspartate residues located 8 and 10 amino acids, respectively, downstream from the first of these two cysteines. This hypothetical signature for neurotransmitter receptor-channel complexes²⁴ is conserved only poorly in the protein encoded by GluR-K1. Although the proline and aspartate residues are present, the first cysteine residue is located only 7 residues upstream from the proline residue, and the second cysteine residue is absent (Fig. 5a).

A hydropathy plot analysis of GluR-K1 (Fig. 4) reveals several regions that are candidates for transmembrane domains (TMDs). The region between amino acids 455 and 810 is of particular interest as its hydropathy profile resembles that seen in the other ligand-gated ion channels: three closely spaced putative TMDs which are separated by ~175 amino-acid residues from a fourth putative TMD which is located close to the C terminus of the protein. Although the presumed TMD IV

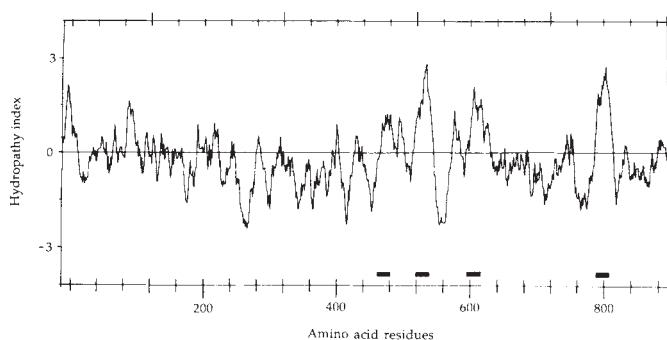


FIG. 4 Hydropathy plot of GluR-K1 (residues +1 to 889), using the algorithm of ref. 44 with a window setting of 15. The calculation was done with the University of Wisconsin software package⁴⁵. The bars mark the regions of the four proposed TMDs.

is very prominent, the assignment of three TMDs in the 'TMD cluster' region (amino acids 521–612) is less obvious. Clearly, there are two very good candidate regions for TMDs, located at amino-acid residues 521–539 and 596–614, respectively. A third TMD might be assigned to the region between amino-acid residues 463 and 480, or alternatively between residues 567 and 585. We favour the interpretation that the region between residues 463 and 480 actually represents a third TMD (see below). Thus, we would assign the following four transmembrane regions in GluR-K1: TMD I, located between amino-acid residues 463 and 480, TMD II between residues 521 and 539, TMD III between residues 596 and 614, and TMD IV between residues 788 and 808. The region between residues 463 and 480 is proposed as TMD I for two reasons. First, in all ligand-gated ion channels the putative TMD I contains a conserved proline residue^{23,24} which is located at position 12 in all nAChRs, and at position 8 in the GABA and glycine receptors. In the putative TMD I of GluR-K1, a proline is found at position 12. Second, putative TMD II in GluR-K1 shows sequence homology with putative TMD II regions in the nAChRs (Fig. 5b). Homology in this region might be expected in view of the fact that both nicotine-gated as well as glutamate-gated ion channels have to provide the environment for the passage of cations, and TMD II is the region that is thought to line the ion channel of the nAChR³⁰. Taken together, the structural features found in the GluR-K1 protein are consistent with this receptor being a member of the superfamily of ligand-gated ion channels. The homologies detected, however, are not strong enough for an unequivocal conclusion to be reached; this will require the analysis of the primary structures of more members of the glutamate receptor family.

Regional tissue distribution of GluR-K1 RNA

Northern blot analysis with the full-length GluR-K1 cDNA probe (Fig. 6) revealed the presence of large amounts of a RNA species, 5.2 kb in size, in poly(A)⁺ RNA isolated from cortex, cerebellum and hippocampus, and trace amounts of RNA in the brainstem. Additionally, two minor bands corresponding to RNAs of sizes ~3.9 and ~3.2 kb were seen in these tissues.

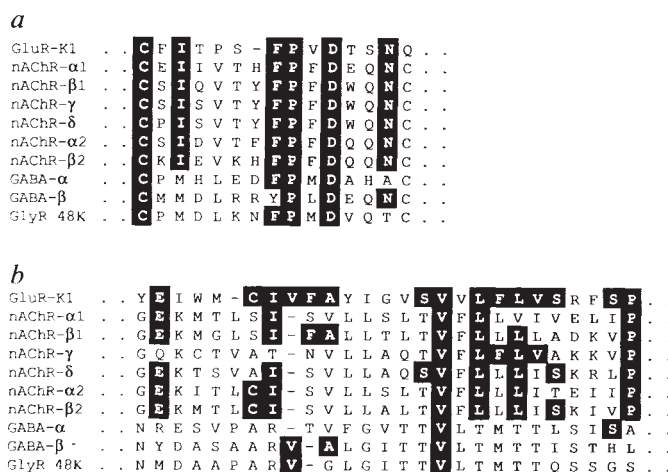


FIG. 5 **a**, Comparison of the extracellular region located between amino-acid residues 89 and 106 of GluR-K1 with the 'Cys-Cys-loop' region found in all other ligand-gated ion channels, showing sequence homology. Sequences of nAChR subunits α , β 1, γ , δ are from mouse muscle⁴⁶, those of nAChR subunits α 2 and β 2 are from rat brain²⁶. GABA_A subunits α and β are from calf brain²⁴. GlyR 48K is the 48K subunit of the glycine receptor from rat brain²⁹. Boxed amino acid residues are found at identical positions in GluR-K1 as well as in at least one of the other receptor sequences. One gap has been introduced arbitrarily. **b**, Comparison of the putative TMD II region of GluR-K1 with hypothetical TMD II regions of other ligand-gated ion channels^{24,26,29,46} (as **a**), suggesting protein-sequence conservation.

These minor bands were also observed under conditions of high-stringency hybridization (see legend to Fig. 6), ruling out the possibility of cross-hybridization to distantly related mRNAs. Most probably, these minor bands represent GluR-K1 RNA transcripts emanating from different polyadenylation sites. Control tissue from outside the CNS (liver and lung) showed no traces of GluR-K1 RNA, and neither did poly(A)⁻ RNA from any of the tissues tested. The regional distribution of GluR-K1 RNA in the brain correlates well with autoradiographic studies of kainate-binding sites, which showed the highest density to be in the hippocampus, with lesser density in cortex and cerebellum^{7,31} and undetectable labelling in brainstem.

Discussion

We have isolated a cDNA clone, GluR-K1, from a rat forebrain cDNA library by screening for expression of kainate-gated ion channels in *Xenopus* oocytes. This cDNA encodes a functional glutamate receptor consisting of a single protein of calculated M_r 99,769, disregarding any possible post-translational modifications. This protein forms an ion channel which possesses electrophysiological and pharmacological properties consistent with those reported for kainate receptors studied in more complex systems^{16,18} or cultured neurons³². The size of

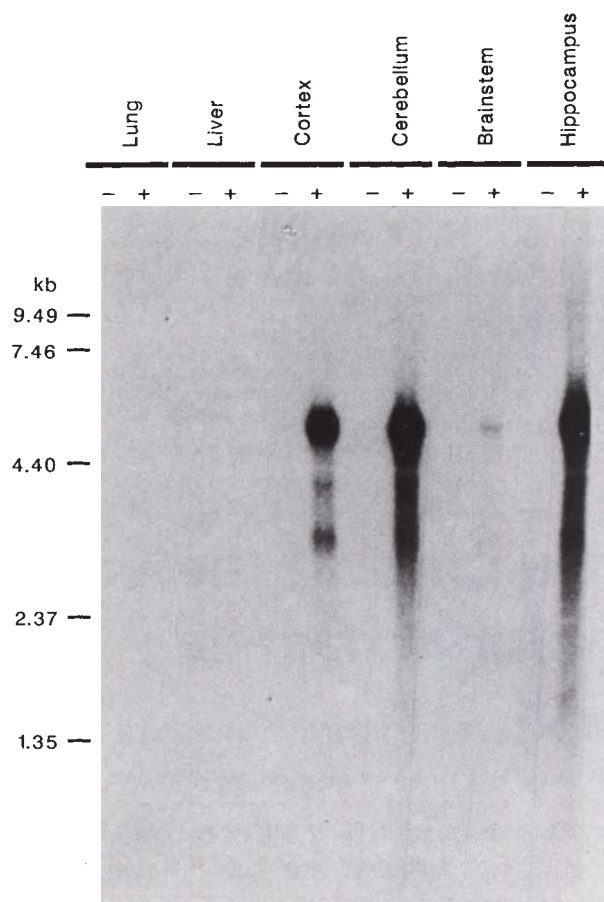


FIG. 6 Northern blot analysis of six rat tissues for the presence of GluR-K1 message. Poly(A)⁺ RNA (3 μ g per lane) and poly(A)⁻ RNA (6 μ g per lane) were prepared as in Fig. 1, separated on 1% formaldehyde-agarose gels and transferred to a Nytran filter. A full-length GluR-K1 DNA probe was prepared by the random prime method using the Amersham Multiprime DNA labelling kit, and was hybridized (10^6 c.p.m. ml⁻¹) to the filters. Prehybridization and hybridization were at 42 °C in 4x SSC/40% formamide, and filters were washed in 0.2x SSC/0.1% SDS at 55 °C, followed by 0.1x SSC/0.1% SDS at 65 °C. The RNA ladder from BRL was used to standardize RNA sizes. + and -, Poly(A)⁺ RNA and poly(A)⁻ RNA, respectively.

the protein encoded by GluR-K1 is an unexpected feature. In fact, none of the ligand-gated ion channel subunits characterized at the molecular level so far has a molecular mass nearly as large as GluR-K1, the largest one previously reported³³ being the nAChR subunit $\alpha 4$ (M_r 67K). Antibodies against a kainate binding protein of M_r 49K from chick cerebellum have been reported to recognize a protein of M_r 92.5K in chick brain tissue, leading to the speculation that both these subunits might be part of a functional kainate receptor¹¹. Furthermore, antibodies recognizing a 48 K kainate binding protein of unknown biological function isolated from frog brain cross-react with a protein of M_r 99K in rat brain preparations³⁴. This cross-reacting component was detected in rat cortex, cerebellum and hippocampus, but not in brainstem or liver, thus matching the distribution of GluR-K1 RNA, as revealed by northern blot analysis (see Fig. 6).

Given the fact that single subunits of many other ligand-gated ion channels have been found to be sufficient for functional expression^{33,35-37}, the finding that GluR-K1 alone is capable of

forming a functional receptor ion channel complex is not surprising. Presumably, the functional receptor-ion-channel complex formed by GluR-K1 is a homo-oligomer. Kainate receptors found *in vivo* could, however, be either homo-oligomers, or hetero-oligomers consisting of GluR-K1 plus additional subunits. In fact, we have isolated several additional cDNA clones which have significant sequence identity to GluR-K1, suggesting the existence of a gene family (J. Boulter and M. H., unpublished observations). Reports that some antibodies to the frog brain kainate-binding protein cross-react with rat brain proteins of M_r 42K, 55K and 57K lend further support to this hypothesis³⁴. Additional subunits, though obviously not required for channel function in oocytes, might be required to convey tissue specificity to kainate receptors, or confer susceptibility to regulatory mechanisms. The latter mechanism has been reported to underlie differences in benzodiazepine pharmacology of GABA_A receptors, which are regulated by three different, highly homologous α -subunits³⁸. □

Received 13 October; accepted 30 October 1989.

1. Cajal, S. R. *Histologie du Système Nerveux de l'Homme et des Vertébrés* (Moloiné, Paris, 1911).
2. Hebb, D. O. *The Organization of Behavior* (Wiley, New York, 1949).
3. Monaghan, D. T., Bridges, R. J. & Cotman, C. W. *A. Rev. Pharmacol. Toxicol.* **29**, 365-402 (1989).
4. Collingridge, G. L. & Bliss, T. V. P. *Trends Neurosci.* **10**, 288-293 (1987).
5. Foster, A. & Fagg, G. E. *Brain Res. Rev.* **7**, 103-164 (1984).
6. Sugiyama, H., Ito, I. & Hirano, C. *Neurosci. Lett.* **325**, 531-533 (1987).
7. Monaghan, D. T., Holets, V. R., Toy, D. W. & Cotman, C. W. *Nature* **306**, 176-179 (1983).
8. McKernan, R. M., Castro, S., Poat, J. A. & Wong, E. H. F. *J. Neurochem.* **52**, 777-785 (1989).
9. Henley, J. M. & Barnard, E. A. *J. Neurochem.* **53**, 140-148 (1989).
10. Hampson, D. R. & Wentholt, R. J. *J. Biol. Chem.* **263**, 2500-2505 (1988).
11. Gregor, P., Eshhar, N., Ortega, A. & Teichberg, V. I. *EMBO J.* **7**, 2673-2679 (1988).
12. Klein, A. U., Niederöst, B., Winterhalter, K. H., Cuenod, M. & Streit, P. *Neurosci. Lett.* **95**, 359-364 (1988).
13. Chen, J.-W., Cunningham, M. D., Galton, N. & Michaelis, E. K. *J. Biol. Chem.* **263**, 417-426 (1989).
14. Brose, N., Halpain, S., Suchanek, C. & Jahn, R. *J. Biol. Chem.* **264**, 9619-9625 (1989).
15. Masu, Y. *et al. Nature* **329**, 836-838 (1987).
16. Hirano, C., Ito, I., Yamagishi, S. & Sugiyama, H. *Neurosci. Res.* **6**, 106-114 (1988).
17. Gundersen, C. B., Miledi, R. & Parker, I. *Proc. R. Soc. B* **221**, 127-143 (1984).
18. Lerma, J., Kushner, L., Spray, D. C., Bennett, M. V. L. & Zukin, R. S. *Proc. natn. Acad. Sci. U.S.A.* **86**, 2083-2087 (1989).
19. Verdoorn, T. A. & Dingleline, R. *Molec. Pharmacol.* **34**, 298-307 (1988).
20. Davies, J. & Watkins, J. C. *Brain Res.* **206**, 172-177 (1981).
21. Fagg, G. E. *Trends Neurosci.* **8**, 207-210 (1985).
22. von Heijne, G. *Nucleic Acids Res.* **14**, 4683-4690 (1986).
23. Grenningloh, G. *et al. Nature* **330**, 25-26 (1987).
24. Barnard, E. A., Darlison, M. G. & Seeburg, P. *Trends Neurosci.* **10**, 502-509 (1987).
25. Feng, D. F., Johnson, M. S. & Doolittle, R. F. *J. molec. Evol.* **21**, 112-125 (1985).
26. Deneris, E. S., Boulter, J., Swanson, L. W., Patrick, J. & Heinemann, S. *J. J. Biol. Chem.* **264**, 6268-6272 (1989).
27. Duvoisin, R., Deneris, E. S., Patrick, J. & Heinemann, S. *Neuron* **3**, 487-496 (1989).

28. Schofield, P. R. *et al. Nature* **328**, 221-227 (1987).
29. Grenningloh, G. *et al. Nature* **328**, 215-220 (1987).
30. Giraudat, J., Dennis, M., Heidmann, T., Chang, J. Y. & Changeaux, J. P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2719-2723 (1986).
31. Patel, S., Meldrum, B. S. & Collins, J. F. *Neurosci. Lett.* **70**, 301-307 (1986).
32. Perouansky, M. & Grantyn, R. *J. Neurosci.* **9**, 70-80 (1989).
33. Boulter, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 7763-7767 (1987).
34. Hampson, D. R., Wheaton, K. D., Dechesne, C. J. & Wentholt, R. J. *J. Biol. Chem.* **264**, 13329-13335 (1989).
35. Blair, L. A. C., Levitan, E. S., Marshall, J., Dionne, V. E. & Barnard, E. A. *Science* **242**, 577-579 (1988).
36. Shivers, B. D. *et al. Neuron* **3**, 327-337 (1989).
37. Schmieden, V., Grenningloh, G., Schofield, P. R. & Betz, H. *EMBO J.* **8**, 695-700 (1989).
38. Pritchett, D. B., Lüddens, H. & Seeburg, P. H. *Science* **245**, 1389-1392 (1989).
39. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochem.* **18**, 5294-5299 (1979).
40. Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
41. Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. *Nucleic Acids Res.* **16**, 7583-7600 (1988).
42. Messing, J., Gronenborn, B., Müller-Hill, B. & Hofschneider, P. H. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3642-3646 (1977).
43. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
44. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105-132 (1982).
45. Devereux, J., Haeblerli, P. & Smithies, O. *Nucleic Acids Res.* **12**, 387-395 (1984).
46. Heinemann, S. *et al. in Molecular Neurobiology: Recombinant DNA Approaches* (eds Heinemann, S. & Patrick, J.) 45-96 (Plenum, New York, 1987).

ACKNOWLEDGEMENTS. This work was supported by the Deutsche Forschungsgemeinschaft (M.H.), by the National Institute of Neurological and Communicative Disorders and Stroke and the Fritz B. Burns Foundation (S.H.), and by an NIH postdoctoral fellowship to S.W.R. We thank John Connolly for introducing us to oocyte recording techniques, and for his suggestion to use voltage recording for the screening procedure. We thank Jim Boulter and Irm Hermans-Borgmeyer for advice and for reading the manuscript, Karin Doebeili for help with the oocyte preparations, Stratagene for making available a pre-release version of their Uni-ZAP™ directional cloning kit, and all our colleagues for helpful discussion.

Structure of the core and central channel of bacterial flagella

Keiichi Namba, Ichiro Yamashita & Ferenc Vonderviszt

ERATO, Molecular Dynamic Assembly Project 5-9-5, Tokodai, Tsukuba 300-26, Japan

X-ray fibre diffraction analysis of bacterial flagellar filaments has allowed the subunit packing and secondary structure arrangement in the filament core to be determined. The central hole, presumably a channel for flagellin transport, is large enough to accommodate the folded elongated flagellin molecules during their transport to the distal end for filament growth.

as a propeller to drive the cell through its liquid environment. The bulk of the flagellum is composed of a single protein, flagellin. Flagellin from *Salmonella typhimurium* wild-type strain SJW1103 consists of 489 amino-acid residues and has a relative molecular mass (M_r) of 51,000 (51K) (ref. 4). The structural design of the flagellar filament is interesting in many respects. First, it is difficult to see how the helical filament can be constructed by bonding identical subunits to each other in a regular manner. Second, flagella adopt several distinct polymorphic forms with different helical parameters in response to chemical and physical changes in their environment such as those of pH, ionic strength or the direction of motor rotation⁵⁻¹⁰. Third, the flagellar filament can self-assemble¹¹. There have been many structural studies on bacterial flagella, and three-dimensional

BACTERIA swim by using one or more flagellar filaments driven by rotary motors¹⁻³. The flagellar filament is helical, and acts

FIG. 1 X-ray diffraction pattern from reconstituted L-type straight flagellar filaments. *a*, Observed diffraction pattern. Straight-type flagellar filaments were isolated from *S. typhimurium* strain SJW1660 (provided by Dr R. Kamiya, Nagoya University) as previously described²⁴. The filaments were depolymerized into monomers and were purified by fast protein liquid chromatography (FPLC) on a Mono-Q column (Pharmacia). The monomers were repolymerized into the filaments with the desired length distribution by adding 1 M ammonium sulphate to a flagellin solution (10 mg protein ml⁻¹). The reconstituted filaments were washed in 0.2 M glycine-NaOH, pH8, and were centrifuged for 20 h at 10,000g. The resulting soft pellet was transferred into a capillary tube. The filaments were orientated by being forced to flow back and forth in the capillary for 15 min to 2 h. The protein concentration of the specimen was 100–150 mg ml⁻¹. We used reconstituted filaments because native filaments are ~10 μ m long and therefore the resulting pellet is too hard to orientate. The filaments were reconstituted from flagellin monomers so that the length distribution could be controlled. Well-oriented specimens were obtained with filaments of an average length of ~0.5 μ m. Specimen preparation will be described in more detail elsewhere (I.Y., F.V. T. Noguchi and K.N., manuscript in preparation). Fibre diffraction experiments were done with a RIGAKU RU200 rotating anode X-ray generator using a 0.1-mm focal cup, operated at 40 kV and 25 mA and double mirror optics. CuK α -radiation (λ = 1.5418 Å) was used. The diffraction patterns were recorded on a photostimulable phosphor-plate (Imaging Plate, Fuji Film) with a 20-h exposure and a specimen-to-film distance of 170 mm. Recorded images were scanned with a Fuji Film BA100 Imaging Plate scanner using 0.1 mm rasters and were displayed on a Seiko D-SCAN GR4416 graphics terminal. Photographic copies were made with a Dunn Instruments MultiColor type 638 colour hardcopy camera. The much higher sensitivity and wider dynamic range of the Imaging Plate compared with those of photographic films³³ made the search for well-oriented specimens efficient and made it possible to record diffraction patterns with high signal-to-noise ratios in a relatively short time. *b*, Simulated diffraction pattern. The layer-line intensities were extracted from the diffraction pattern shown in *a* with the angular deconvolution method^{25,26} and the diffraction pattern was reconstructed using these intensities and the angular spread function without background.

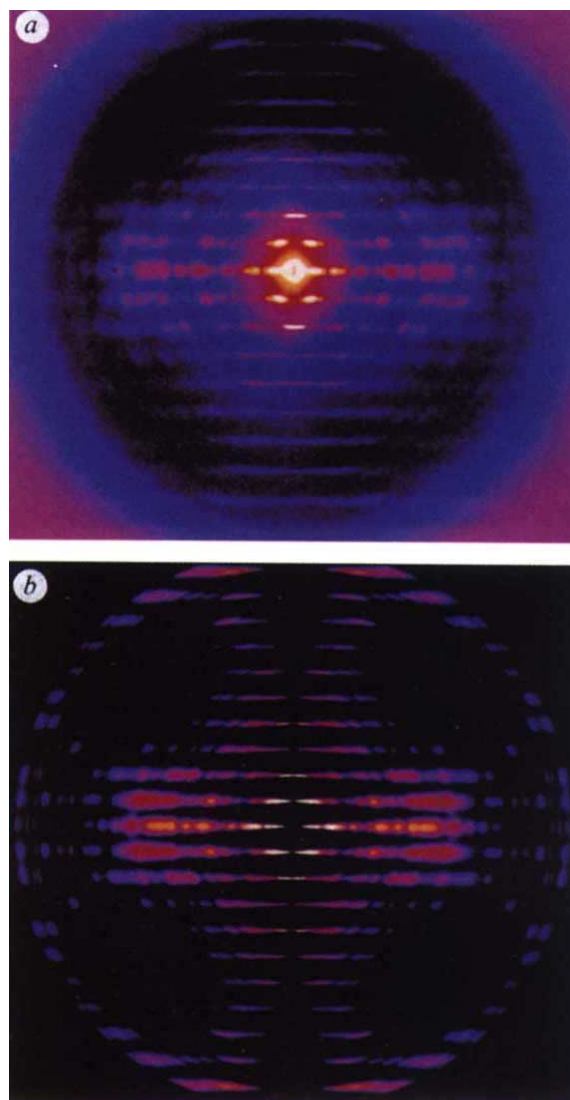


image reconstructions from electron micrographs have shown the subunit shape on the surface of the filament^{12,13}. It has not yet been possible, however, to visualize the subunit packing in the filament to explore the architecture of this assembly. Also, flagellar growth *in vivo* is known to occur at the distal end of the filament^{14,15}, where the HAP2 protein forms a cap structure, which presumably serves to prevent arriving flagellin subunits from diffusing away before polymerizing^{16,17}. It is thought that the flagellin molecules are transported through the central channel of the filament, but the channel structure has not yet been unambiguously determined.

The helical form of the wild-type filament is postulated to involve two different conformations of flagellin monomers in one filament^{5,18–21}. The structural symmetry, therefore, is complicated and analysis of wild-type filaments is not feasible. Straight filaments, in which all subunits are in the same conformation, would make structural analysis possible. Such filaments have been isolated from mutants of *S. typhimurium*^{22,23}.

We have solved the structure of a straight-type mutant bacterial flagellum at a resolution of 20 Å, using X-ray fibre diffraction. The electron density map clearly shows the shape of the whole subunit as well as the packing and indicates that the innermost domain of each subunit is mostly responsible for filament formation. The map also shows the central channel of the filament, which is large enough for folded flagellin monomers to pass through. Based on other evidence, we have made tentative assignments of primary sequence to the morphological

domains observed. In particular, because a proteolytic fragment of flagellin missing portions of the terminal regions cannot polymerize into the filament²⁴, the innermost domain was assigned to the amino- and carboxy-terminal segments. The intensity distribution at 10-Å resolution in the diffraction pattern indicates that the innermost domain, which forms the core of the filament, consists of a bundle of α -helices aligned in parallel to the filament axis.

X-ray fibre diffraction

There are two types of straight filament (L- and R-type), whose subunits are assembled with opposite-handed helical symmetry. Figure 1*a* shows a fibre diffraction pattern from reconstituted L-type straight filaments from *S. typhimurium* strain SJW1660. Flagellin from this strain differs from that of the wild-type strain SJW1103 by only one amino acid (S. Kanto and S.-I. Aizawa, personal communication). Because the diffraction pattern shows very little interference from lateral lattice formation, it is equivalent to the cylindrically averaged diffraction pattern from a single filament. The helical symmetry is reasonably well approximated by 331 subunits in 60 turns of the basic helix. The pitch of the basic helix is 26.35 Å, as measured from layer lines 2, 4 and 6, where the Bessel orders 1, 2 and 3 contribute, respectively. The axial rise per subunit is 4.78 Å, as measured from layer line 11, where the Bessel order 0 contributes. The structure repeats every 1,581 Å along the filament. Layer lines indexed by 1,581 Å spacing are not individually visible because

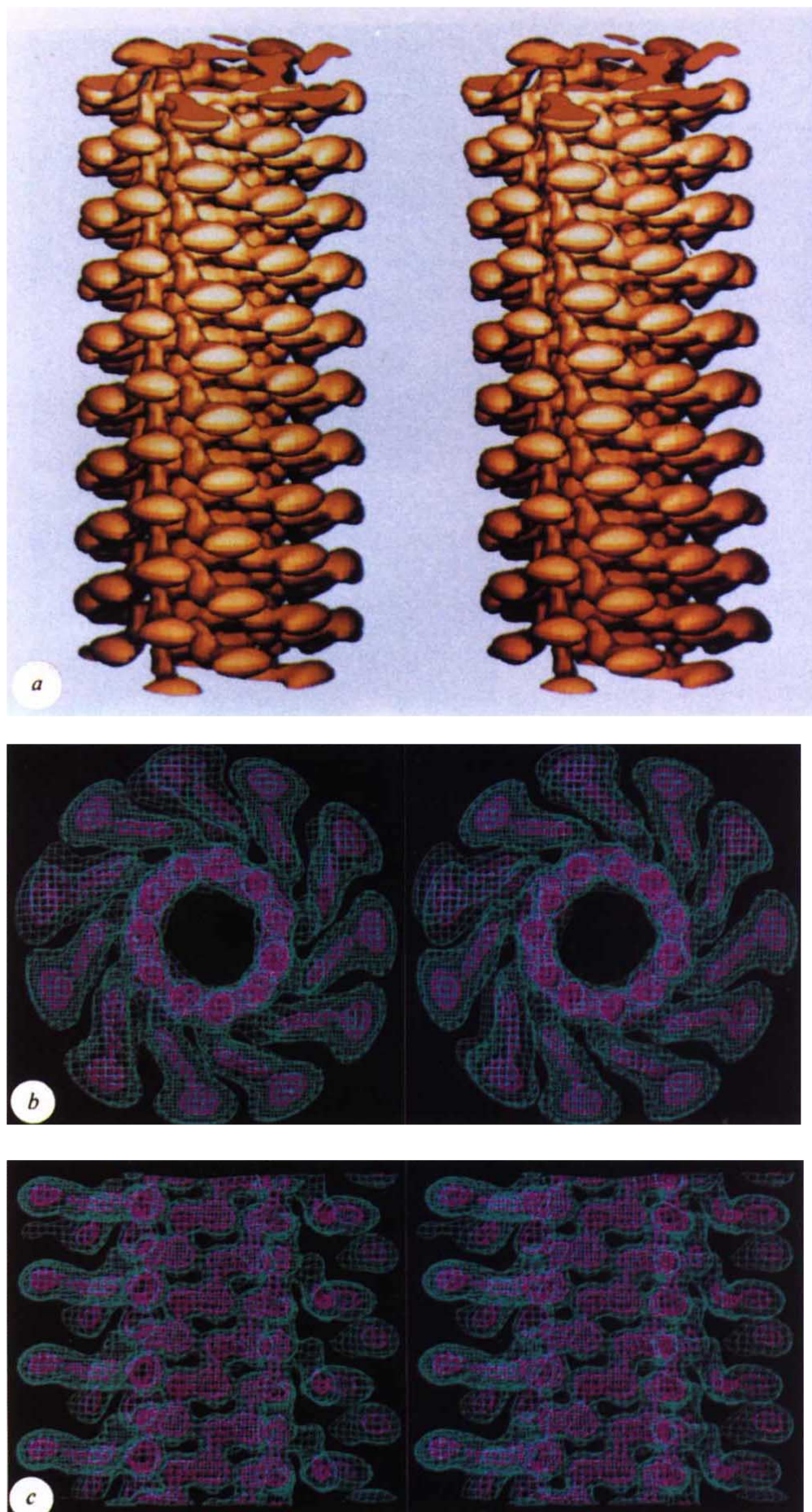


FIG. 5 Refined electron density maps. *a*, Oblique view of a shaded solid representation of the contoured map. The map is contoured at level 30, which gives $\sim 80\%$ of the correct volume. The length of the map along filament axis is 500 Å. *b*, Cross-section of the filament with 52.5 Å axial thickness. *c*, Longitudinal section. The section is 55 Å thick and 200 Å long axially. The position of this section is indicated on the left side of *a*. The maps in *b* and *c* are both contoured at two levels of 30 and 90.

of disorientation within the sample and the finite X-ray beam size. Many layer lines are fused to make up each layer line in Fig. 1a. In Fig. 2, a surface lattice shows the helical array of subunits and a corresponding n, l plot describes the diffraction pattern. The apparent layer line 0 in Fig. 1a, for instance, contains (layer line, Bessel order): 0, 0; 1, -11; -1, 11; 2, -22; -2, 22; and so on.

The diffraction pattern in Fig. 1a was processed by the angular deconvolution method^{25,26} to extract intensities along layer lines. The angular width of disorientation estimated by this process was 3°. A diffraction pattern after background subtraction is simulated in Fig. 1b, using the layer-line intensities and the angular spread function deduced from the diffraction pattern. The general intensity distribution clearly shows the strong intensity in the equatorial region around 10 Å resolution. This indicates the presence of α -helices aligned with their axes almost parallel to the filament axis, which is also seen in other systems such as α -keratin, myosin and fibrinogen^{27,28}. The maximum dimension of an object diffracting X rays can be estimated from the spatial frequency of the diffraction pattern if a continuous transform is available. The radial position of the α -helices was estimated to be up to ~70 Å from the layer-line intensity distributions.

Structure analysis

Three-dimensional image reconstruction from electron micrographs has been done by Trachtenberg and DeRosier on frozen hydrated preparations of native flagellar filaments isolated from strain SJW1660 (ref. 13). We used the phases from the image reconstruction and the amplitudes collected from the X-ray fibre diffraction pattern to make a starting model for phase refinement. Figure 3 shows a comparison between their data and our X-ray data of the square roots of the layer-line intensities. The square roots are quite similar, which indicates the high quality of the images in the electron micrographs of the frozen hydrated specimens. There are notable differences, however, partly because of the contrast transfer function in electron microscopy and partly because of differences in sample preparation. The electron micrographs were of native filaments¹³, whereas the X-ray fibre diffraction patterns were from reconstituted filaments.

Maps of cross-sections in different stages of phase refinement are shown in Fig. 4. The sections are 52.5-Å thick and contoured at a relatively low (Fig. 4a, c, e) or a high level (Fig. 4b, d, f). The phases used in three-dimensional reconstruction¹³ shown in Fig. 4a and b were combined with amplitudes measured by X-ray fibre diffraction, and an electron density map was calculated (Fig. 4c, d). A molecular envelope was calculated on the basis of this map and the solvent-flattening phase refinement procedure was carried out 10 times²⁶. This procedure is as follows. The solvent region in an initial map is flattened and the phases from the modified map calculated. These phases are then combined with observed amplitudes, and an electron density map is calculated. The resulting map is shown in Fig. 4e and f.

The initial map derived from three-dimensional image reconstruction¹³ clearly shows a well-defined shape for the subunits on the filament surface, but the density of the map of the inner core is weak and noisy (Fig. 4a, b). The density in the core region became reasonably high when phases were combined with the X-ray amplitudes (Fig. 4c, d). After 10 cycles of phase

refinement, the subunit boundaries became much clearer (Fig. 4e, f). The electron densities are continuous and evenly distributed throughout the structure. Contoured maps at various levels of density show very stable features, which is an indication of good phasing.

Subunit shape and packing

The overall surface structure of the filament is shown in Fig. 5a as a shaded solid model. The arrangement of the subunits can be compared with the surface lattice in Fig. 2a. The inner part of the filament shows strong connections between subunits, forming a dense core of 30–55 Å radius, whereas there are many extensive spaces and gaps in the outer part. The subunit density extends ~15 Å from the core, goes down ~25 Å and extends out again horizontally, skewed along the azimuth about 60° from the radius, to an outer radius of 120 Å.

Figure 5b shows a stereo-image of a cross-section of the filament. It is viewed down the filament axis from the cell. The section is 52.5-Å thick and contains 11 subunits in two turns of the basic helix. The basic helix goes through every other subunit in this view. Neighbouring subunits are staggered ~26 Å along the direction of view. The molecular boundaries of the subunits are clear enough to show the subunit shape in most parts. Although the subunits are closely packed in the core of the filament and it is difficult to determine the subunit boundaries, there is a strong density fluctuation with 11-fold symmetry in this axial view, representing the 11 nearly longitudinal protofilaments, each of which is assumed to be a cooperative unit in the two-state model for polymorphism^{5,18–21}.

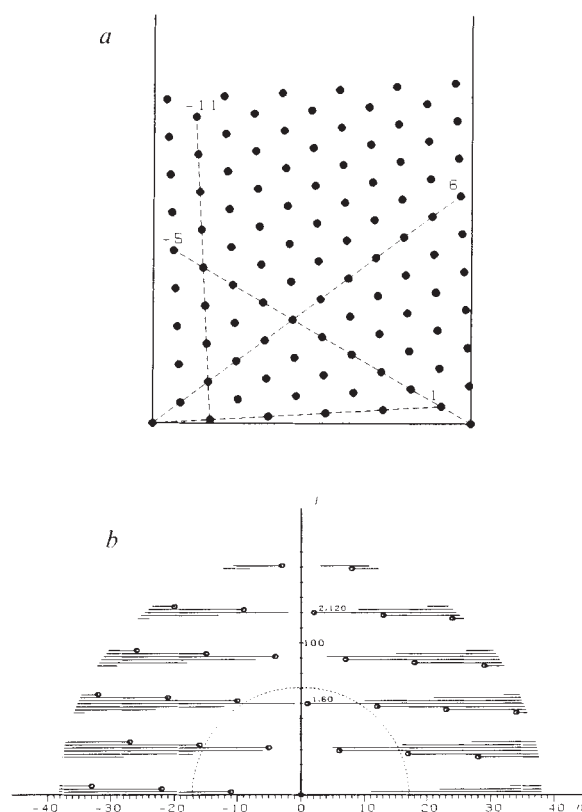


FIG. 2 A surface lattice and an n, l plot. a, Surface lattice showing the helical arrangement of subunits in the L-type straight flagellar filament³⁴. Some of the helical families are marked with dashed lines. b, Plot (n, l) describing the diffraction pattern in Fig. 1a. n , Order of the Bessel functions; l Layer-line numbers. The resolution limit used in the structural analysis is marked with a dotted line.

TABLE 1 M_s s of domains in the flagellin subunit based on volume fractions estimated from the contoured electron density map

Domain	1	2	3
Volume fraction	0.3	0.25	0.45
M_r	15K	13K	23K

Figure 5c is a longitudinal section of the filament viewed from inside, being 55-Å thick and 200-Å long along the axis; the region comprising this section is indicated on the left side of Fig. 5b. Axial stagger of the subunits is visible in the middle of the section shown in Fig. 5c. Neighbouring subunit relations are along the 5-start left-handed helix to the top right, along the 6-start right-handed helix to the top left, and along the basic helix nearly horizontally as indicated in Fig. 2a. The domain contacts are so close in the core that it is hard to distinguish from the lower contour whether the vertical rod density lying along the 11-start helix or the horizontal density along the basic helix belongs to a single subunit. From the higher contour the latter is more likely to belong to a single subunit. Then, the subunit contacts are extensive along the basic and the 5-start helices, and less extensive along the 6-start helix.

A subunit can be divided into three domains, which are labelled in Fig. 6a. Measuring along the filament radius, domains 1, 2 and 3 extend from 30–55 Å, 55–80 Å and 80–120 Å, respectively. The innermost domain (domain 1) has the highest density and is responsible for most of the filament formation. The middle domain (domain 2) seems to have some interaction with domain 3 of the neighbouring subunit in the clockwise direction, but this interaction is not as close as that between domains 1 of neighbouring subunits. Domain 3 is on the filament surface and is presumed to determine the antigenic properties. Domain 1 seems to consist of two globular subdomains, only one of which is extensively connected with domain 2 within a subunit. Domain 2 is a vertical rod of 35-Å length, slightly tilted towards the filament axis. Domain 3 appears somewhat fin-shaped, 50 Å long and 50 Å wide at its widest part with an axial thickness of ~25 Å.

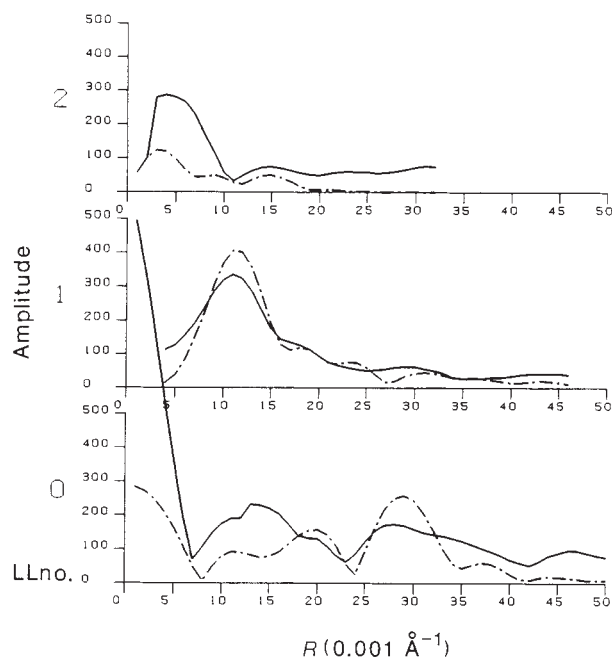


FIG. 3 A comparison of layer-line amplitudes. Solid line, X-ray fibre diffraction data; dashed line, data used in three-dimensional image reconstruction from electron micrographs¹³. The data represented by the dashed line were calculated by grouping layer-line intensities from the original data¹³ according to the n/l plot in Fig. 2b. It was not possible to obtain the data for the peak at the origin of layer line 0 from the X-ray diffraction pattern because of the interfilament-interference effect observed here. Therefore, the data from the image reconstruction work¹³ were scaled-up three times so that the initially negative radial density distribution at the filament axis became zero. Adjusting the negative density level to zero improves the molecular envelope, resulting in the better-refined images. LL no., Layer-line number.

Assignment of amino-acid sequence

The volume of each domain was estimated from the refined map and the corresponding molecular weights are listed in Table 1. The distribution of α -helical secondary structure in flagellin was predicted from the primary sequence and is plotted in Fig. 6b, with the domain boundary positions determined by proteolytic digestion²⁴. F40 and F27 represent fragments with M_s of 40K and 25K, respectively, which were isolated from a digestion mixture. The predicted secondary structure is consistent with circular dichroic spectra of flagellin and its fragments in the far ultraviolet region (F. V., H. Uedaira, S. Kidokoro and K.N., manuscript in preparation). It has also been shown that the terminal regions of monomeric flagellin are partially unfolded and flexible²⁴.

Because the terminal regions of flagellins of different serotype or even from different species have high amino-acid homology, these terminal regions are believed to be important in filament formation^{4,29,30}. Indeed, the F40 fragment, which is missing the 65 N-terminal residues and 44 C-terminal residues of flagellin, cannot form filaments even under conditions that favour polymerization²⁴. It is reasonable, therefore, to assign these terminal regions to domain 1, the estimated volume of which is consistent with such an assignment. As the F27 fragment contains the region of highly variable sequence responsible for flagellar antigenicity³¹, it probably corresponds to domain 3 located on the filament surface; the estimated volume of domain 3 is consistent with this. Domain 2 can, therefore, be assigned to the parts of F40 absent from F27 on the basis of both deduction and estimated volume. These assignments are summarized in Fig. 6. A part of D2 indicated in Fig. 6b could, however, be contributing to the filament core, because the amino-acid replacements which

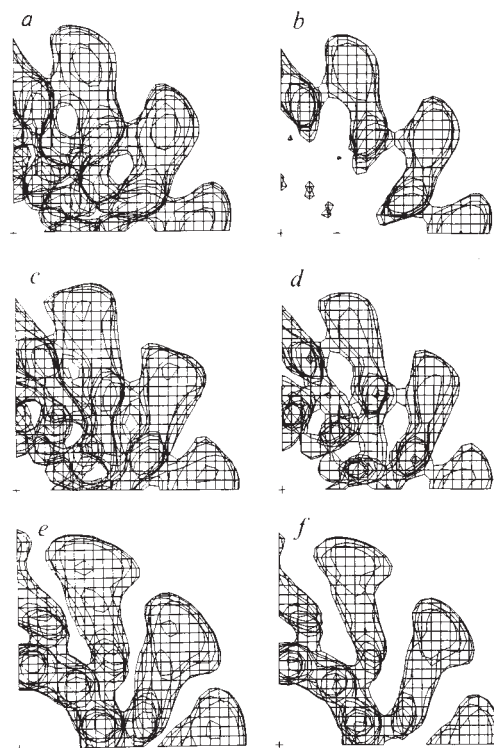
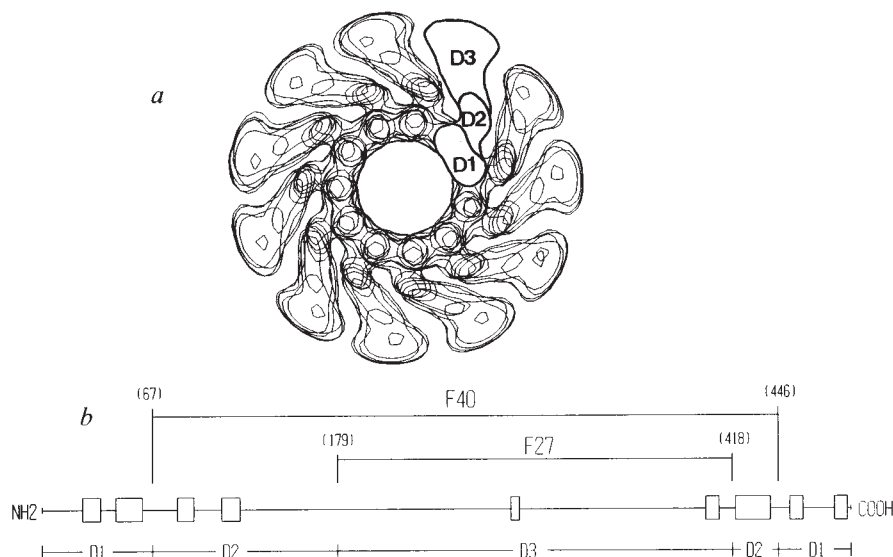


FIG. 4 Improvement of the electron density maps during the phase refinement procedure. Only a quarter of a cross-section of the filament is shown for each map. The sections are 52.5 Å thick. a, b, Three-dimensional image reconstruction from electron micrographs of frozen hydrated specimens¹³. c, d, Maps calculated with the amplitudes from the X-ray diffraction pattern and the phases from the three-dimensional image reconstruction. e, f, Maps calculated after 10 cycles of the solvent-flattening phase refinement procedure. The maps are contoured at density level 30 in a, c and e, and 50 in b, d and f, where the highest density is ~150 in arbitrary units.

FIG. 6 Amino-acid sequence assignment for domains in the electron density map. *a*, Contoured electron density map viewed along the filament axis. Twelve sections at 4.7-Å intervals are superimposed and the domains are labelled. *b*, Diagram showing the predicted distribution of α -helices and the domain boundary positions. Rectangles represent α -helices. The secondary structure was predicted using the program ALB³⁵ obtained from Brookhaven Protein Data Bank. The domain boundary positions were determined by proteolytic digestion²⁴.



affect filament shape are in the D2 region and the interspecific sequence homology of this region is as high as that of D1^{4,29,30}. Thus, the sequences with high α -helical content are assigned to domain 1 and 2 and the non- α -helical part to domain 3; this is consistent with the interpretation of the intensity distribution at ~ 10 -Å resolution in the diffraction pattern. Preliminary model fitting shows that the cross-section of domain 1 shown in Fig. 6*a* is just the right size to consist of two bundles, each containing four α -helices, arranged side by side with their axes parallel to the filament axis. Differential scanning calorimetry of F27 suggests that domain 3 contains two thermodynamically independent units (F.V., H. Uedaira, S. Kidokoro and K.N., manuscript in preparation), although it is not visible in the map at the present resolution.

Central channel

Although it is known that the growth of flagellar filaments *in vivo* occurs at the distal end^{14,15}, it is not known how flagellin monomers reach the tip of the filament. Although the best explanation is that flagellin monomers pass through the filament's central channel, there has not been any experimental evidence showing that the central hole is large enough to accommodate folded flagellin molecules. Unlike tobacco mosaic virus³², electron micrographs of negatively stained filaments do not show the hollow channel, for reasons that are unknown. Three-dimensional image reconstructions from electron micrographs¹³ show a small hole (Fig. 4*a*), which could suggest that flagellin monomers are transported in an unfolded and stretched conformation. But the density distribution of the core-part forming the hole was weak and noisy. The refined electron density map using X-ray fibre diffraction data shows well-defined subunit packing in the core region and a central hole of ~ 60 -Å diameter, which is large enough to accommodate the folded flagellin molecule with the elongated shape shown in Fig. 6*a*. It is probable, therefore, that flagellin molecules are transported through the central channel with the same conformation as they have in solution. A hole of this size should be visible by electron microscopy. Electron micrographs of transverse thin sections of reconstituted filaments show a large central hole, supporting the result described above, whereas native filaments have a filled appearance, indicating that the central hole of native filaments is filled with flagellin molecules (T. Akiba, C. J. Jones and K.N., manuscript in preparation).

Implications for flagellar function

Having made the correspondence, shown in Fig. 6, between the volumes of the domains that appear as distinct entities in Fig. 5 and fragments of flagellin defined in biochemical assays, it is

easy to understand how each domain of the flagellin subunit can be independently responsible for its respective role. Antigenicity is an important function of the highly variable domain 3; this variability can leave the filament structure undisturbed, because domain 1 alone seems to be able to form the filament backbone regardless of the conformation of domain 3, provided it is not oversized. If the boundary of domain 1 shown in Fig. 6*a* is correct, subunit addition during polymerization proceeds along the basic helix at the bottom of the filament shown in Fig. 5*a*, where only one site would be available for subunit addition at a time. About 100 residues at the terminal regions of flagellin, which are in a flexible conformation in solution (ref. 24; and S.-I. Aizawa, F.V., R. Ishima and K. Akasaka, manuscript in preparation), become well folded in an α -helical conformation when bound to the end of the basic helix. Because it is difficult to conceive how these terminal regions could be recognized by the binding site if they were totally unfolded, they must be in a somewhat folded conformation. In agreement with this, circular dichroic spectra indicate the presence of some unstable α -helical structure in these regions of monomeric flagellin²⁴.

The model proposed for flagellar polymorphism^{5,18-21} assumes high cooperativity within a protofilament running along an 11-start helix, and independence between protofilaments. The structure presented here does not directly support this model, because one of the two subunit-contacts along a protofilament is relatively weak and a subunit extends over two protofilaments. But the structure presented here is not inconsistent with the model either, because high cooperativity does not necessarily require a direct connection along the protofilament, and a protofilament consisting of two subdomains of domain 1, each coming from neighbouring subunits, could be independent from neighbouring protofilaments, to which the other subdomains of domain 1 contribute, if subdomains behave as independent subunits. In fact the present structure, in which a subunit appears to extend over two protofilaments, seems to correspond to the scheme proposed by Calladine^{19,20}. Because the resolution of the structure at present is not high enough to permit the confident description of the subunit connections, further attempts to reconcile it with the current model of polymorphism are premature. Comparison of this structure with that of R-type straight filament should help resolve this question. □

Received 15 August; accepted 17 October 1989.

1. Berg, H. C. & Anderson, R. A. *Nature* **245**, 380-382 (1973).
2. Silverman, M. & Simon, M. *Nature* **249**, 73-74 (1974).
3. Larsen, S. H., Reader, R. W., Kort, E. N., Tso, W. W. & Adler, J. *Nature* **249**, 74-77 (1974).
4. Joys, T. M. *J. biol. Chem.* **260**, 15758-15761 (1985).

5. Asakura, S. *Adv. Biophys.* **1**, 99–155 (1970).
6. Kamiya, R. & Asakura, S. *J. molec. Biol.* **106**, 167–186 (1976).
7. Kamiya, R. & Asakura, S. *J. molec. Biol.* **108**, 513–518 (1977).
8. Macnab, R. M. & Ornston, M. K. *J. molec. Biol.* **112**, 1–30 (1977).
9. Kamiya, R., Hotani, H. & Asakura, S. *Symp. Soc. exp. Biol.* **35**, 53–76 (1982).
10. Hotani, H. *J. molec. Biol.* **156**, 791–806 (1982).
11. Asakura, S., Eguchi, G. & Iino, T. *J. molec. Biol.* **10**, 42–56 (1964).
12. Shirakihara, Y. & Wakabayashi, T. *J. molec. Biol.* **131**, 485–507 (1979).
13. Trachtenberg, S. & DeRosier, D. J. *J. molec. Biol.* **195**, 581–601 (1987).
14. Iino, T. *J. gen. Microbiol.* **56**, 227–239 (1969).
15. Emerson, S. W., Tokuyasu, K. & Simon, M. I. *Science* **169**, 190–192 (1970).
16. Ikeda, T., Asakura, S. & Kamiya, R. *J. molec. Biol.* **184**, 735–737 (1985).
17. Ikeda, T., Homma, M., Iino, T., Asakura, S. & Kamiya, R. *J. Bacteriol.* **169**, 1168–1173 (1987).
18. Wakabayashi, K. & Mitsui, T. *J. phys. Soc. Japan* **33**, 175–182 (1972).
19. Calladine, C. R. *Nature* **225**, 121–124 (1975).
20. Calladine, C. R. *J. theor. Biol.* **57**, 469–489 (1976).
21. Calladine, C. R. *J. molec. Biol.* **118**, 457–479 (1978).
22. Kamiya, R., Asakura, S. & Yamaguchi, S. *Nature* **286**, 628–630 (1980).

23. Fujita, H., Yamaguchi, S., Taira, T. & Iino, T. *J. gen. Microbiol.* **125**, 213–216 (1981).
24. Vonderviszt, F., Kanto, S., Aizawa, S.-I. & Namba, K. *J. molec. Biol.* (in the press).
25. Makowski, L. *J. appl. Crystallogr.* **11**, 273–283 (1978).
26. Namba, K. & Stubbs, G. *Acta crystallogr. A* **41**, 252–262 (1985).
27. Astbury, W. T. *Proc. R. Soc. B* **134**, 303–328 (1947).
28. Astbury, W. T. *Proc. R. Soc. B* **141**, 1–9 (1953).
29. Wei, L.-N. & Joys, T. M. *J. molec. Biol.* **186**, 791–803 (1985).
30. Trachtenberg, S. & DeRosier, D. J. *J. molec. Biol.* **202**, 787–808 (1988).
31. Fedorov, O. V., Kostyukova, A. S. & Pyatibratov, M. G. *FEBS Lett.* **241**, 145–148 (1988).
32. Finch, J. T. *J. molec. Biol.* **8**, 872–874 (1964).
33. Amemiya, Y., Satow, Y., Matsushita, T., Chikawa, J., Wakabayashi, K. & Miyahara, J. *Topics Curr. Chem.* **147**, 121–144 (1988).
34. Kamiya, R., Asakura, S., Wakabayashi, K. & Namba, K. *J. molec. Biol.* **131**, 725–742 (1979).
35. Finkelstein, A. V. & Ptitsyn, O. B. *Biopolymers* **22**, 15–25 (1983).

ACKNOWLEDGEMENTS. We thank Drs H. Hotani, F. Oosawa, S.-I. Aizawa and D. L. D. Caspar for continual support and discussions, Drs S. Trachtenberg and D. J. DeRosier for making data available before its publication, and T. Tibbitts for help in preliminary model fitting.

LETTERS TO NATURE

Enhanced shock acceleration in relativistic jets and cosmic ray origin in active galactic nuclei

J. J. Quenby & R. Lieu

Astrophysics Group, Blackett Laboratory, Imperial College, London SW7 2BZ, UK

THE shocked terminations of the relativistic plasma jets emitted by active galactic nuclei (AGNs) have been considered as possible sources of high-energy cosmic rays¹. However, the acceleration rate provides a stringent limitation on the ability of this class of source to make particles with the highest observed energies². Here we report a numerical simulation of the diffusive shock acceleration process when the flow is relativistic. The acceleration rate is found to be ~ 13.5 -times faster than the previous values used, which were based on an analytical approximation. Significant particle anisotropies appear in the process. This enhancement of acceleration allows even 10^{19} -eV protons and 10^{20} -eV iron nuclei to be produced in AGNs of the local supercluster.

Diffusive acceleration of test particles in a parallel shock of planar geometry, when the flow speeds V_1 (upstream), V_2 (downstream) are nonrelativistic and when the test-particle speed v is approximately the speed of light, c , can be treated analytically^{3–6}. The most important results are (1) that the number spectrum of advected particles is

$$n(p) \propto p^{-\alpha} \quad (1)$$

where p is the particle momentum, $\alpha = (r+2)/(r-1)$ and $r = V_1/V_2$ is the shock compression ratio; (2) that the average time necessary to accelerate a particle from initial momentum p_i to final momentum p_f ($p_f \gg p_i$) is

$$t = \frac{c}{V_1 - V_2} \int_{p_i}^{p_f} \left(\frac{\lambda_1}{V_1} + \frac{\lambda_2}{V_2} \right) \frac{dp}{p} \quad (2)$$

where λ_1 and λ_2 are the diffusion mean free paths upstream and downstream respectively (λ/p is usually constant for relativistic particles).

Here we report results from a Monte Carlo routine that simulates test-particle acceleration in parallel shocks with arbitrary upstream flow speeds. A test particle is injected at many mean free paths upstream of the shock and is carried by the fluid until it has drifted far away downstream. The distance x between scattering events and the polar angle θ that the particle velocity vector makes with respect to the shock normal are both random variables, with probability distributions of $\exp(-x/\lambda)/\lambda$ and $\sin \theta$ respectively. The latter procedure simulates isotropic scattering, which is a viable alternative model to pitch-angle diffusion in strong-turbulence situations.

For nonrelativistic flow speeds ($V_1, V_2 \ll c$), the numerical simulation reproduces diffusion theory, that is equations (1) and (2) and other details, such as $\Delta p/p$ (defined as the fractional momentum increase per cycle—or per pair of shock crossings encounters), the escape probability per cycle and angular dependence in particle distribution at the shock.

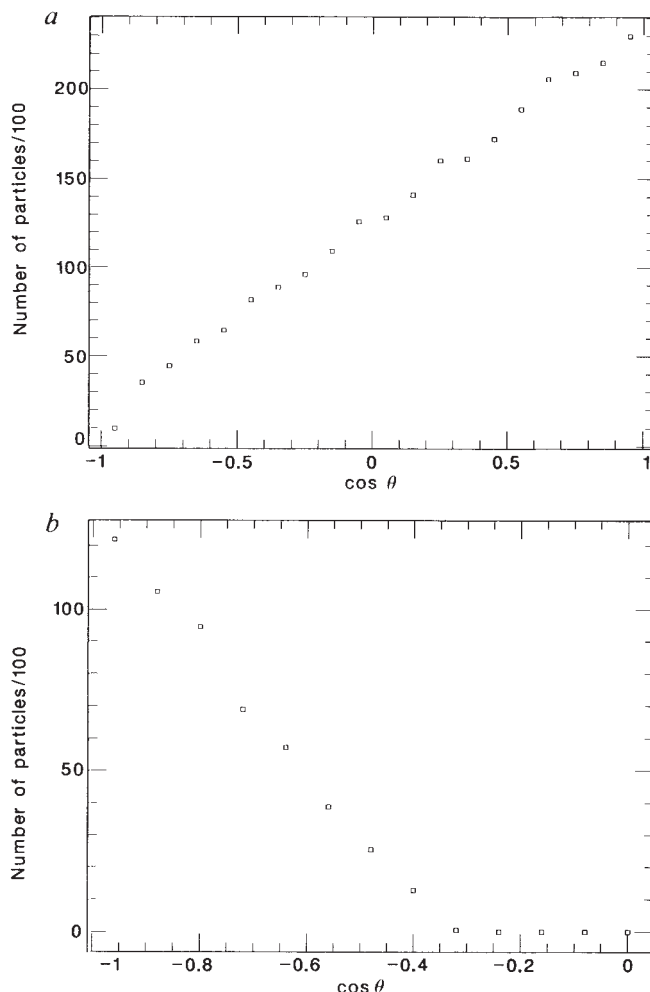


FIG. 1 Frequency distribution of $w = \cos \theta$ for particles that cross the shock from upstream (a) and downstream (b), as obtained by Monte Carlo simulation. The angle that the particle velocity vector makes with the shock normal, θ is measured in the respective fluid frames. The shock under consideration has $V_1 = 0.96c$, $V_2 = 0.32c$. The characteristic maxima at $w = 1$ and $w = -1$ are also found in Kirk and Schneider²¹ (see their Fig. 6 for the case of large-angle scattering). If the shock were nonrelativistic, the corresponding distributions would have been directly proportional to w and $-w$.

If the upstream flow speed becomes relativistic ($V_1 \leq c$), anisotropies in the particle distribution at the shock invalidate the diffusion approach (see Fig. 1). Numerical simulation, on the other hand, shows that $\Delta p/p \gg 1$, with a resulting enhancement in the acceleration rate.

For $V_1 = 0.96c$, $V_2 = 0.32c$ (that is, for a relativistic compression ratio r of 3, see ref. 7), Fig. 2 shows the frequency distribution of $\Delta p/p$, which averages at ≈ 10.5 . The large resultant fractional energy increase is, on average, approximately γ^2 , where $\gamma = (1 - V^2/c^2)^{-1/2}$, and V is the relativistic difference between V_1 and V_2 . The average relativistic cycle time t_r , for a given initial momentum p_i , is also provided by the model. Typically, within a time t_r , a cosmic-ray particle in the shock vicinity crosses it twice, and is accelerated from momentum p_i to momentum $p_f = (1 + \Delta p/p)p_i \approx 11.5p_i$. To compare the relativistic result with diffusion theory, the nonrelativistic acceleration time, t_{nr} given by equation (2), is calculated using identical values of p_i , p_f , V and λ/p . It is found that $t_{nr}/t_r \approx 13.5$.

Figure 3 shows how an escape probability per cycle P can be deduced from a frequency histogram giving the number of cycles before escape. Because neither this, nor $\Delta p/p$ are small compared with unity, the particle number power-law index of $-1 - (P/(\Delta p/p))$ does not develop in the initial few decades of energy beyond the injection energy.

Relativistic shocks have been studied by Bogdan and Webb⁸, and by Kirk and Schneider⁹. The former used a two-stream approximation, wherein particle velocity directions are reversed in each scattering. Their results are more dramatic than ours. The latter, on the other hand, used pitch-angle diffusion with a small mean scattering angle, and reported considerably milder acceleration. Owing to the breakdown of diffusion theory in relativistic shocks, a sequence of many small-angle scatterers can no longer be replaced by a few large-angle scatterers with longer mean free path. Because our results correspond to isotropic scattering, they fall somewhere in between those of these two previous works.

To support our use of isotropic scattering centres, we develop an argument based on scaling known properties of energetic-particle scattering in the interplanetary medium. Cosmic-ray modulation studies indicate that $\lambda \propto L_m$ (magnetic rigidity) in the relativistic regime¹⁰. Detailed trajectory integration in a field model based on the region around an interplanetary shock reveal large pitch-angle changes ($\Delta \cos \theta \sim 0.5-1$) occurring in times that are short compared with the travel time between these events (see Figs 2a, b, c of ref. 11). The mean free path λ derived for

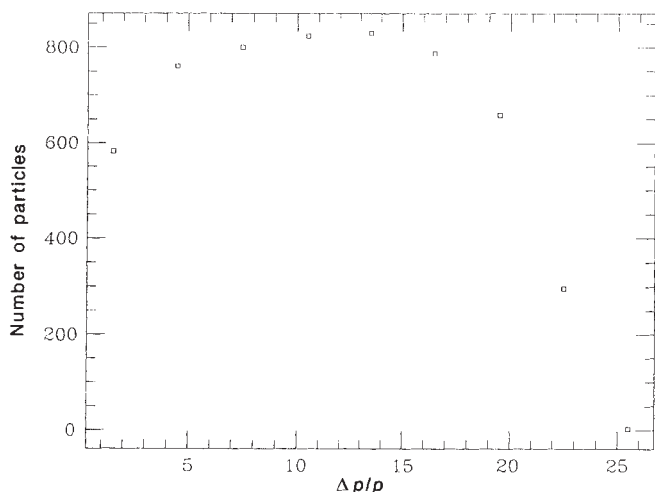


FIG. 2 Frequency distribution of $\Delta p/p$, the fractional change of cosmic-ray momentum within a cycle of downstream-upstream-downstream traversals, for a parallel relativistic shock with $V_1 = 0.96c$, $V_2 = 0.32c$.

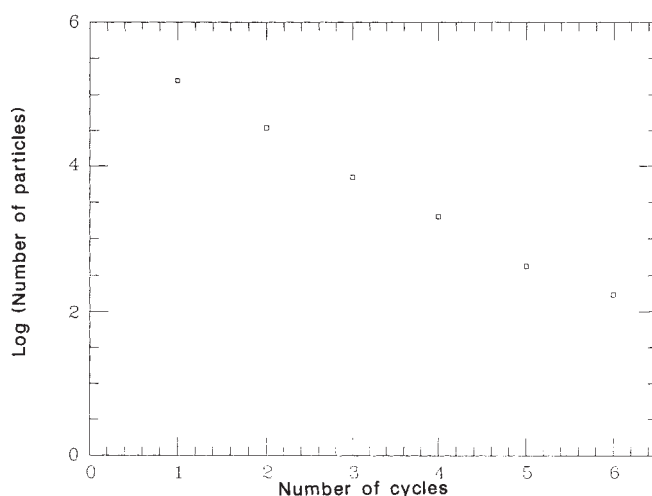


FIG. 3 Frequency distribution of n , the number of downstream-upstream-downstream cycles that a cosmic-ray particle makes before it leaves the shock, for a parallel shock with $V_1 = 0.96c$, $V_2 = 0.32c$. Note that the frequency is proportional to $(1 - \alpha)^n$, where α is the escape probability per cycle. As the frequency is given on a logarithmic scale, the plot yields an approximate straight line of slope $-\alpha$.

this particular data set can be fitted by $\lambda = K\rho_c$ for a particle cyclotron radius ρ_c with the empirical constant $K = 41$ (ref. 12). For $L_m = 200$ MV we have $\lambda = 0.03$ AU which is roughly the separation seen between the large-amplitude waves existing behind the shock wave we are studying (see Figs 1a, b of ref. 13). Quasilinear scattering theory, based on a series of small pitch-angle changes, certainly does not apply in the disturbed ($\Delta B/B \approx 0.3$) interplanetary field¹⁰, and we suggest that such knowledge as we have of AGN jets and shocks indicates a similar regime.

Magnetic fields in AGN hotspots, the shock working surface of our model, are $\sim 4 \times 10^{-4}$ G. Hence for large-angle scattering of particles with momentum $10^{16}-10^{19}$ eV/c (where c is the speed of light) we need structures separated by corresponding λ -values that are 1 pc-1 kpc. But large-scale inhomogeneities in synchrotron emission are seen¹⁴ in the jet of 3C120, for example, on scales of ~ 10 mas-10 arcsec, corresponding approximately to these λ -values. Whether these structures are individual small-scale shocks or instabilities or other nonlinear wave effects, they must all feed equivalent-scale inhomogeneities into the hotspot shock at the jet termination.

Although observation of relativistic jet speeds is only convincing near the core and according to some calculations there is a significant slowing before the hotspot, the data are still consistent with maintenance of relativistic speeds throughout¹⁴, and one theory involves re-acceleration of a plasma¹⁵. Observation of one-sided jet sources reveal more depolarization on the opposite (non-jet) side¹⁶. This depolarization is continuous in angular dependence and, in one case at least, arises from a foreground effect. Doppler beaming as a reason for one-sidedness and relativistic flow on all length scales are likely explanations¹⁷. Other parameters for the shock working surface are taken from Laing¹⁸ who found field strengths of $\sim 4 \times 10^{-4}$ G at the hotspot and dimensions of this region of ~ 20 kpc from a survey of AGN jets. The field tended to be parallel to the jet, before and within the hot spot, suggesting a parallel-shock model.

For particles of rigidity 10^{19} V accelerated by a planar parallel shock, $V_1 = 0.96c$, $V_2 = 0.32c$, $\lambda = 41\rho_c$ with a 4×10^{-4} G mean field, the equation (2) acceleration time constant, enhanced by the calculated factor of 13.5 is $\approx 2 \times 10^4$ yr. This is short compared with the proton synchrotron and 2.7-K photoionization loss times of 9×10^7 yr and 10^{10} yr respectively. The diffusive loss

time, T_d from the medium over a distance R of 10 kpc is 2×10^5 yr according to $T_d \approx 3R^2/4\lambda c$, which is still longer than the acceleration time. At 10^{20} V, however, $\lambda = 10$ kpc according to our model, so diffusive acceleration is not likely at this rigidity. Iron nuclei at rigidities $< 10^{19}$ V, however, can supply the energy for the 10^{19} – 10^{20} -eV particles.

Following Ginzburg^{19,20}, the origin of 10^{17} – 10^{20} -eV cosmic rays is due to acceleration within the local supercluster (~ 30 Mpc in size) in which conservatively ten AGNs may be described by the parameters adopted here. An individual AGN power output of $2.6 \times 10^{51}/T$ erg s⁻¹ is required to explain the observed $> 10^{17}$ -eV spectrum for a loss time T in the supercluster. Taking an intercluster field of 10^{-8} G yields a diffusive loss time of 2×10^9 yr at 10^{17} eV, and nuclear collision and photoionization losses are negligible. The required jet power into $> 10^{17}$ -eV particles of 5×10^{34} erg s⁻¹ is small compared to that seen in the radiofrequency range ($\leq 10^{46}$ erg s⁻¹) and even if the accelerated, low-energy cosmic-ray gas significantly loads the jet because of wave-particle interaction, the high end of the spectrum can still be treated by the test-particle approach.

The actual spectrum of $> 10^{17}$ -eV cosmic-ray protons expected for our galaxy on the basis of the above model depends on the source spectrum, the momentum-dependent diffusion in the supercluster and the spectrum of source sizes. For a single source at the centre of a homogeneous diffusing medium with a free-escape boundary at R and a catastrophic particle-loss time constant T_e , the propagation equation

$$\frac{\lambda c}{3} \nabla^2 n(p, r) - \frac{n}{T_e} = -Q\delta(r) \quad (3)$$

has as a leading term in the $r \ll R$ solution, $n \propto \lambda^{-1}$ (ref. 19). Here, Q is the source function, r is the distance from the centre of the diffusing medium and $\delta(r)$ is the Dirac function. Thus the dominance of diffusive loss in the supercluster is expected to steepen the calculated differential-source spectrum, which has a p^{-1} dependence, by one power of p , because $\lambda \propto p$. If, in addition, the upper cutoff rigidity of each of the AGN hotspot contributions increases with available jet energy, there is a further spectral steepening. The 10^{17} – 10^{20} -eV cosmic-ray spectrum is believed to be $n \propto p^{-3}$, although a heavy-nuclei origin of the highest-energy air showers would require a further steepening. Some contribution to the overall spectrum below the knee in the primary cosmic-ray spectrum is expected, but this is not important if the escape probability from the galaxy reduces to a $p^{-0.5}$ dependence and the energetic-particle loading of the accelerator eventually sets in. We note that, experimentally, this feature may extend as low as 10^{13} eV. $\square$

Does mass accretion lead to field decay in neutron stars?

N. Shibazaki*, T. Murakami†, J. Shaham‡
& K. Nomoto§

* Department of Physics, Rikkyo University, Tokyo 171, Japan

† Institute of Space and Astronautical Science, Sagami-hara, Kanagawa 229, Japan

‡ Department of Physics and Columbia Astrophysics Laboratory, Columbia University, New York, New York 10027, USA

§ Department of Astronomy, Faculty of Science, University of Tokyo, Tokyo 113, Japan

WHETHER or not neutron-star magnetic fields decay is a matter of current debate. The recent observation^{1,2} of cyclotron lines from γ -ray-burst sources, thought to be relatively old neutron stars, indicates that they are strongly magnetized and therefore that their fields have not decayed. One interpretation of the correlation observed³ between the strength of the magnetic field and the mass accreted by the neutron star is that mass accretion may itself lead to the decay of the magnetic field. Adopting the hypothesis of accretion-induced field decay, we calculate here the consequent evolution of a neutron star's spin and magnetic field. The results are consistent with observations of binary and millisecond radio pulsars. Thermomagnetic effects⁴ could provide a possible physical mechanism for such accretion-induced field decay.

Statistical analyses^{5,6} of ~ 400 radio pulsars and the study³ of the origin and evolution of magnetized neutron stars in binary systems are consistent with the assumption that neutron stars are born with magnetic fields of 10^{12} G which then decay with a time constant of $5 - 10 \times 10^6$ yr. Estimates of age and magnetic field strength of binary millisecond radio pulsars, however, require⁷ that the field decay, if it occurs, should stop or proceed much more slowly on timescales of $\geq 10^9$ yr, at field strengths $\leq 10^{10}$ G. Pulsar models in which magnetic fields do not decay but align with the rotation axis are shown to also be consistent with the observed properties of radio pulsars⁸⁻¹⁰.

Ohmic dissipation of electrical currents in the crust has been thought to be the physical cause for field decay¹¹. Recent calculations¹² of the ohmic decay of dipolar magnetic fields, however, have shown that the field does not decay exponentially as has been assumed in most statistical analyses of radio pulsars and that, if the field occupies the entire crust, it decays by less than a factor of 100 in a Hubble time.

If the strong magnetic fields of neutron stars do not decay, then the origin of the weak magnetic fields of less than 10^{10} G, found in binary and millisecond radio pulsars¹³, remains to be explained. Neutron stars in binary and millisecond radio pulsars are thought to be formed by the iron-core collapse of massive stars or by the accretion-induced collapse of massive white dwarfs (for reviews see refs 14, 15). In one model for the weak field and rapid rotation of millisecond radio pulsars it is argued that it is simply white dwarfs¹⁶⁻¹⁸ or iron cores of massive stars¹⁹ having the appropriate field strength and angular momentum that give birth to the observed millisecond radio pulsars after a collapse. In another model, mass accretion in low-mass X-ray binaries, which are assumed to be progenitors of millisecond radio pulsars, is proposed to cause both the field decay²⁰ and the spin-up²¹. In fact, Taam and van den Heuvel³ found a possible inverse correlation between the magnetic field strength and the estimated total mass of accreted matter for binary X-ray sources and for binary and millisecond radio pulsars. This inverse correlation supports the latter model. A study of the evolution of magnetic fields in the crust of a neutron star⁴ showed further that the inward heat flux caused by mass accretion gives rise to thermomagnetic effects that could remove the strong magnetic field of a neutron star. Here we examine further the possibility that magnetic fields of neutron stars decay only when

Received 21 June; accepted 3 November 1989.

1. Begelman, M. C., Blandford, R. D. & Rees, M. J. *Rev. Mod. Phys.* **56**, 255–351 (1984).
2. Hillas, A. M. *Astr. Astrophys.* **22**, 425–444 (1984).
3. Krinsky, G. F. *Dokl. Akad. Nauk. SSSR* **234**, 1306–1309 (1977).
4. Axford, W. I., Leer, E. & Skadron, G. *Proc. 15th Int. Conf. on Cosmic Rays* Vol. 11 (ed. Christou, C. Ya.) 132–137 (Bulgarian Acad. Sci., Sofia, 1977).
5. Bell, A. R. *Mon. Not. R. astr. Soc.* **182**, 147–156 (1978); **182**, 443–455 (1978).
6. Blandford, R. D. & Ostriker, J. P. *Astrophys. J.* **211**, 793–808 (1977).
7. Peacock, J. A. *Mon. Not. R. astr. Soc.* **196**, 135–152 (1981).
8. Bogdan, T. J. & Webb, G. M. *Mon. Not. R. astr. Soc.* **229**, 41–52 (1987).
9. Kirk, J. G. & Schneider, P. *Astrophys. J.* **322**, 256–265 (1987).
10. Quenby, J. J. *Space Sci. Rev.* **37**, 201–267 (1984).
11. Moussas, X., Quenby, J. J. & Valdes-Galicia, J. F. *Solar Phys.* **112**, 365–382 (1987).
12. Moussas, X., Quenby, J. J. & Valdes-Galicia, J. F. *Astrophys. Space Sci.* **86**, 185–195 (1982).
13. Mavromichalaki, H. et al. *Solar Phys.* **116**, 377–390 (1988).
14. Walker, R. C., Benson, J. M. & Unwin, S. C. *Astrophys. J.* **316**, 546–572 (1987).
15. Sol, H., Pelletier, G. & Asseo, E. *Mon. Not. R. astr. Soc.* **237**, 411–429 (1989).
16. Laing, R. A. *Mon. Not. R. astr. Soc.* **195**, 261–324 (1981).
17. Garrington, S. T., Leahy, J. P., Conway, R. G. & Laing, R. A. *Nature* **331**, 147–149 (1988).
18. Laing, R. A. *Nature* **331**, 149–151 (1988).
19. Ginzburg, V. L. *Proc. 20th Int. Conf. on Cosmic Rays* Vol. 7 (eds Kozyrsky, V. A. et al.) 7–54 (Nauka, Moscow, 1987).
20. Ginzburg, V. L. & Syrovatskii, S. *The Origin of Cosmic Rays* (Pergamon, Oxford, 1964).
21. Kirk, J. G. & Schneider, P. *Astr. Astrophys.* **201**, 177–188 (1988).

ACKNOWLEDGEMENTS. We thank I. Axford, P. Schneider and J. Kirk for discussions.

neutron stars undergo mass accretion by considering γ -ray bursts and the periods and field strengths of binary and millisecond radio pulsars.

The recent discovery^{1,2} of cyclotron absorption lines with 20-keV and 40-keV energies in γ -ray bursts revealed that the central objects of γ -ray bursters are, as indicated in some theoretical works^{22,23}, strongly magnetized neutron stars with magnetic fields of 10^{12} G. The statistical arguments indicate $\sim 10^7$ yr (refs 24, 25) or $\gg 10^7$ yr (ref. 26) for the age of neutron stars in γ -ray bursters depending on the assumed distance. When combined with the high field strength, the latter age estimate contradicts the field-decay hypothesis whereas the former does not.

We note that the absorption lines were seen only in a limited portion of the burst¹. This can also be used to test the field-decay hypothesis because the presence of the cyclotron lines may depend sensitively on the configuration of the magnetic field relative to the line of sight²⁷. If this is the case, the 5–10-s duration of the cyclotron-absorption feature indicates the rotation period of ≥ 10 –20 s. This rotation period and the field strength put the γ -ray burst sources into the category of the 'turned-off' pulsars⁶. If the spin-down of the neutron star is caused by the magnetic field, which decays exponentially on a timescale of τ_B , the variation of the rotation period P with time t is represented by $P^2 = AB_0^2\tau_B[1 - \exp(-2t/\tau_B)] + P_0^2$ where B_0 and P_0 are initial field strength and rotation period, respectively, and the constant A is 9.8×10^{-40} s G⁻² for a stellar moment of inertia of 10^{45} g cm² and radius of 10^6 cm. To attain the rotation

period of ≥ 10 –20 s from the initial value of ≤ 1 s, magnetic braking requires the initial field strength of the neutron star to be larger than $\sim 2 \times 10^{13}$ G if $\tau_B \sim 10^7$ yr. This value of the initial field seems large, compared to the average values of $\sim 10^{12}$ G indicated by the statistical analyses of radio pulsars. Therefore we pursue here the other possibility—that the magnetic fields of γ -ray burst sources, which might be isolated neutron stars, do not decay.

As already mentioned, the weak magnetic fields of $\leq 10^{10}$ G are found in binary and millisecond radio pulsars¹³; these magnetic fields are plotted against rotation periods in Fig. 1. In the resurrected-pulsar model²¹ the progenitors of millisecond radio pulsars are the low-mass X-ray binaries and their rapid rotations are the consequence of the spin-up that is due to the mass accretion in binary systems. We consider the weak magnetic field and the rapid rotation of these radio pulsars in terms of the hypothesis that mass accretion leads to both spin-up and field decay. We assume that the magnetic field decays with accretion as

$$B = \frac{B_0}{1 + \Delta M/m_B} = \frac{B_0}{1 + \dot{M}t/m_B} \quad (1)$$

where B is the field strength at time t , ΔM the accreted mass, m_B the mass constant for the field decay and $\dot{M}$ the accretion rate. Equation (1) fits well the inverse correlation between the magnetic field strength and the estimated mass of the total accreted matter for binary X-ray sources and binary and millisecond radio pulsars³. As already noted³, this inverse correlation is also consistent with the simple field-decay hypothesis if the larger amount of accreted matter is interpreted simply to reflect the greater age of the neutron star. Although we do not advocate any physical model leading to equation (1), we do assume that a direct relation between field loss and accreted mass exists. We note that equation (1) represents a change in the whole field, not just the component perpendicular to the rotation axis. Using the formula given by Ghosh and Lamb²⁸ for the accretion torque, the variation of rotation period is described by $\dot{P} = -0.11 I^{-1} (GM)^{3/7} (BR^3)^{2/7} n \dot{M}^{6/7} P^2$ where n is the dimensionless accretion torque (see ref. 28), R the radius, I the moment of inertia and G the gravitational constant. The calculated evolutionary tracks in the plot of the magnetic field against rotation period are shown by the solid lines in Fig. 1. The rotation periods and magnetic fields decrease with time. We ended our calculations when the inner edge of the accretion disk reached the surface of the neutron star. When the mass constant for the field decay is $m_B \geq 10^{-3} M_\odot$, the evolution proceeds along the equilibrium rotation line²⁸ (dash-dotted line in Fig. 1), where the spin-up torque that is due to accreting matter and the spin-down torque that is due to magnetic field are balanced. This is because the timescale of the field decay is longer than the spin-up timescale. After mass accretion stops, the parameters develop horizontally towards the right in Fig. 1. If $m_B \geq 10^{-4} M_\odot$, the calculated evolutionary tracks are consistent with B and P of binary and millisecond radio pulsars. These evolutionary tracks are also approximately consistent with the B and P of low-mass binary X-ray sources such as suggested from the beat-frequency model²⁹ for the quasiperiodic X-ray oscillations (see ref. 30 for a review). Note that the possible inverse correlation³ between B and ΔM is represented well by equation (1) with $m_B \sim 10^{-4} M_\odot$. Hence, if we adopt $m_B \sim 10^{-4} M_\odot$, the accretion-induced field-decay hypothesis represented by equation (1) is consistent with B , P and ΔM as observed or estimated for binary and millisecond radio pulsars and binary X-ray sources.

The magnetic fields of single radio pulsars are in the range of 10^{11} – 10^{13} G with a distribution peak around 10^{12} G. The accreted mass of interstellar matter onto a single radio pulsar is negligible compared to $m_B \sim 10^{-4} M_\odot$. Hence, according to the accretion-induced field-decay hypothesis, the present field

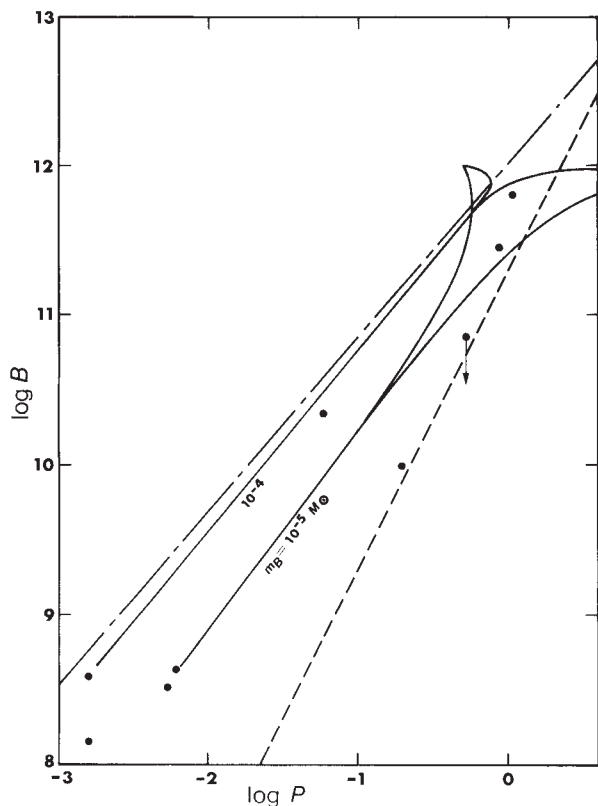


FIG. 1 Evolutionary tracks (solid lines) of neutron stars in the plot of magnetic field as a function of rotation period. Rotation periods and magnetic fields become smaller with accreted mass and hence with increasing time. After mass accretion stops, the parameters develop horizontally to the right. The initial magnetic field is taken as $B_0 = 10^{12}$ G, the initial rotation period is chosen as $P_0 = 0.5$ or 100 s, and the mass accretion rate is fixed at $\dot{M} = 1.1 \times 10^{18}$ g s⁻¹. The mass, radius and moment of inertia of the neutron star are taken as $1.4 M_\odot$, 106 cm, and 10^{45} g cm², respectively. The dash-dotted and dashed lines denote the equilibrium rotation²⁸ and pulsar-death lines⁶, respectively. The filled circles represent the positions of binary and millisecond radio pulsars in this diagram.

strengths of single radio pulsars should be equal to their initial values. But this assertion conflicts with the results of the statistical analyses^{5,6} of radio pulsars, which suggest that surface fields decay on a timescale of $5-10 \times 10^6$ yr. Pulsar statistics depend, however, on the assumptions made for the radio luminosity law, the braking index, the magnetic-field evolution, the distribution of initial field strength and so on. To resolve the above conflict, the statistical analysis under the assumption of no field decay should be conducted, examining especially the dependence on the assumptions for the above properties.

Mass accretion of neutron stars is likely to give rise to the inward heat flow through the crust. The thermomagnetic effects

in the crust that arise from this inward heat flux, such as suggested in ref. 4, have been invoked as a possible mechanism of the accretion-induced field decay. The formula of ref. 4 predicts, in its simplest form, a stronger dependence of the field decay on time than does equation (1). The field-decay formula based on thermomagnetic effects, however, may depend on the structure of the crust, the accretion process and the history of the heat flux. A more detailed study of the thermomagnetic effects in the crust is needed to resolve the controversial issue of the field decay in neutron stars. Other possible mechanisms for the accretion-induced field decay should also be explored. □

Received 18 May; accepted 26 October 1989.

1. Murakami, T. *et al.* *Nature* **335**, 234-235 (1988).
2. Fenimore, E. E. *et al.* *Astrophys. J.* **335**, L71-L74 (1988).
3. Taam, R. E. & van den Heuvel, E. P. J. *Astrophys. J.* **305**, 235-245 (1986).
4. Blondin, J. M. & Freese, K. *Nature* **323**, 786-788 (1986).
5. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613-639 (1985).
6. Stollman, G. M. *Astr. Astrophys.* **178**, 143-152 (1987).
7. Kulkarni, S. R. *Astrophys. J.* **306**, L85-L89 (1986).
8. Candy, B. N. & Blair, D. G. *Astrophys. J.* **307**, 535-539 (1986).
9. Cheng, A. F. *Astrophys. J.* **337**, 803-813 (1989).
10. Kundt, W. *Comments Astrophys.* **12**, 113-121 (1988).
11. Ostriker, J. P. & Gunn, J. E. *Astrophys. J.* **157**, 1395-1417 (1969).
12. Sang, Y. & Chanmugan, G. *Astrophys. J.* **323**, L61-L64 (1987).
13. Kulkarni, S. R. in *Physics of Neutron Stars and Black Holes* (ed. Tanaka, Y.) 37-53 (Universal Academy, 1988).
14. Van den Heuvel, E. P. J. in *The Origin and Evolution Neutron Stars* (eds Helfand, D. J. & Huang, J.-H.) 393-406 (Reidel, Dordrecht, 1987).
15. Nomoto, K. in *The Origin and Evolution Neutron Stars* (eds Helfand, D. J. & Huang, J.-H.) 282-303 (Reidel, Dordrecht, 1987).
16. Michel, F. C. *Nature* **329**, 310-312 (1987).

17. Chanmugan, G. & Brecher, K. *Nature* **329**, 696-698 (1987).
18. Grindlay, J. E. & Bailyn, C. D. *Nature* **336**, 48-50 (1988).
19. Brecher, K. & Chanmugan, G. *Nature* **338**, 712 (1989).
20. Van den Heuvel, E. P. J., van Paradijs, J. A. & Taam, R. E. *Nature* **322**, 153-155 (1986).
21. Alpar, M. A., Cheng, A. F., Ruderman, M. A. & Shaham, J. *Nature* **300**, 728-730 (1982).
22. Michel, F. C. *Astrophys. J.* **290**, 721-727 (1985).
23. Liang, E. *Comments Astrophys.* **12**, 35-49 (1987).
24. Van Paradijs, J. A. *Mon. Not. R. astr. Soc.* **238**, 45-48 (1989).
25. Hartmann, D., Epstein, R. I. & Woosley, S. E. *Astrophys. J.* (in the press).
26. Wang, J. C. L. *et al.* *Phys. Rev. Lett.* (submitted).
27. Bussard, R. W. & Lamb, F. K. in *Gamma-Ray Transients and Related Astrophysical Phenomena* (eds Lingenfelter, R. E., Hudson, H. S. & Worrall, D. M.) 189-200 (American Institute of Physics, 1982).
28. Ghosh, P. & Lamb, F. K. *Astrophys. J.* **234**, 296-316 (1979).
29. Alpar, M. A. & Shaham, J. *Nature* **316**, 239-241 (1985).
30. Lewin, W. H. G., van Paradijs, J. & van der Klis, M. *Space Sci. Rev.* **46**, 273-377 (1988).

ACKNOWLEDGEMENTS. N.S. and T.M. thank J.A. van Paradijs for discussions. J.S. thanks Y. Tanaka for hospitality. This research was supported by the Ministry of Education, Science and Culture (Japan) and NASA.

A hierarchy of catastrophes as a succession of stability limits for the crystalline state

J. L. Tallon

Interdisciplinary Research Centre in Superconductivity, Madingley Road, Cambridge CB3 0HE, UK

MANY theories of melting have focused on some form of instability or catastrophe delimiting the range of stability of the crystalline state. These are exemplified by the vibrational instability proposed by Lindemann¹ and the elastic shear instability of Born². Fecht and Johnson³ suggested that an entropy catastrophe delimits the stability range of both a superheated crystal and a supercooled liquid. Drawing on Kauzmann's idea⁴ that the glass transition pre-empt a catastrophe point at which the entropy of supercooled liquid equals that of the crystal, they suggest that a second catastrophe point occurs when the entropy of the superheated crystal equals that of the liquid phase, and that this is pre-empted by melting. Here I argue that these entropy catastrophes are an outer bound on the stability limits of the crystalline and liquid phases, and that a hierarchy of inner catastrophes occurs at reduced superheatings or supercoolings.

Kauzmann showed that, with progressive supercooling, the entropy of a liquid tends towards the entropy of the crystalline state. Were they to cross at some temperature T_g , one would be faced with Kauzmann's paradox that, at temperatures below T_g , the entropy of the ordered crystal would exceed that of the disordered liquid. This situation is prevented by the occurrence of the glass transition. Although the glass transition may be loosely termed a 'kinetic transition', reflecting the fact that the structural relaxation time becomes so long that fluidity is lost, it is seen from Kauzmann's ideas that there is an underlying instability. The kinetic character means that the observed transi-

tion point depends on the cooling rate but, in the limit of infinitely slow cooling, the transition point must remain greater than or equal to the instability point. Moreover, one consequence of the free-volume theory of the glass transition is the prediction of an underlying first- or second-order transition⁵. Thus, there appears a succession of catastrophe barriers to supercooling, first kinetic (corresponding to the glass transition), then thermodynamic and finally entropic (corresponding to the Kauzmann point). If one barrier were to be surmounted, the next would provide the natural ultimate limit for further supercooling. Here I show that further barriers may be deduced and that a similar hierarchy of catastrophes occurs in relation to superheating of

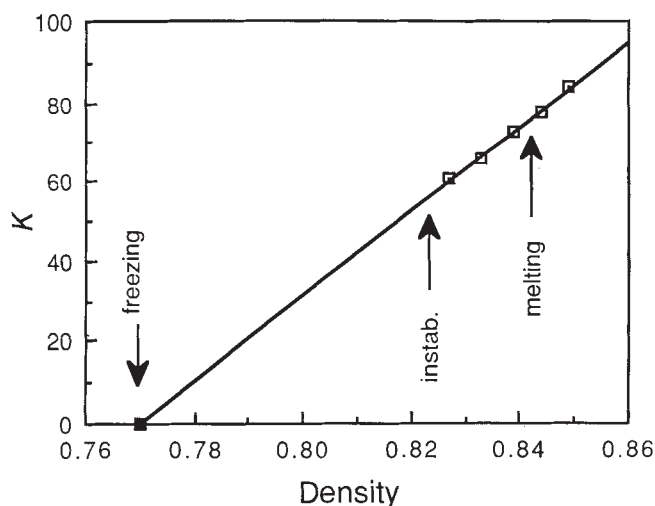


FIG. 1 The Kosterlitz-Thouless parameter K as a function of reduced density for a computer-simulated two-dimensional crystal of Lennard-Jones atoms. The solid data point shows that the crystalline data extrapolate to zero at the freezing density. Densities are quoted in reduced units. The equilibrium freezing and melting points are indicated together with the shear-instability point for maximum superheating.

Permanent address: Physics and Engineering Laboratory, DSIR, PO Box 31313, Lower Hutt, New Zealand.

crystals, that is, to melting and the metastable extent of the crystalline phase.

The interpretation of melting in terms of a one-phase instability model carries with it a number of problems. Born's idea² that the shear modulus falls with increasing temperature becoming zero at the melting point T_m is evidently incorrect, or at least incomplete, as crystals still possess rigidity close to T_m . Mechanical instability models imply second-order melting with marked pre-melting behaviour⁶ rather than the observed first-order melting with little or no pre-transitional behaviour⁷. Finally, the notion of crystalline instability refers only to the properties of the crystalline phase and any rigorous treatment of melting must also refer to the properties of the liquid phase. The melting point is defined as the temperature and pressure at which the Gibbs free energy of the two phases are equal and beyond which thermodynamic instability occurs. Nonetheless the notions of thermodynamic instability and mechanical instability can be harmonized and the two shown to coincide in a special sense. To develop these ideas I use two previously reported results.

First, if one examines the shear or rigidity moduli as functions of molar volume for a large range of cubic or close-packed solids, one of the moduli extrapolates to zero at the volume of the liquid at the freezing point^{8,9}. To illustrate this I have examined computer-simulation data for a two-dimensional Lennard-Jones system and find exactly the same result as previously reported^{8,9} for real three-dimensional systems. The data are the results of Monte Carlo calculations¹⁰ of the elastic Lamé coefficients μ and λ for a 529-atom system close to the triple-point. The Kosterlitz-Thouless parameter K defined by

$$K = \frac{4a^2}{kT_m} \frac{\mu(\mu + \lambda)}{(2\mu + \lambda)} \quad (1)$$

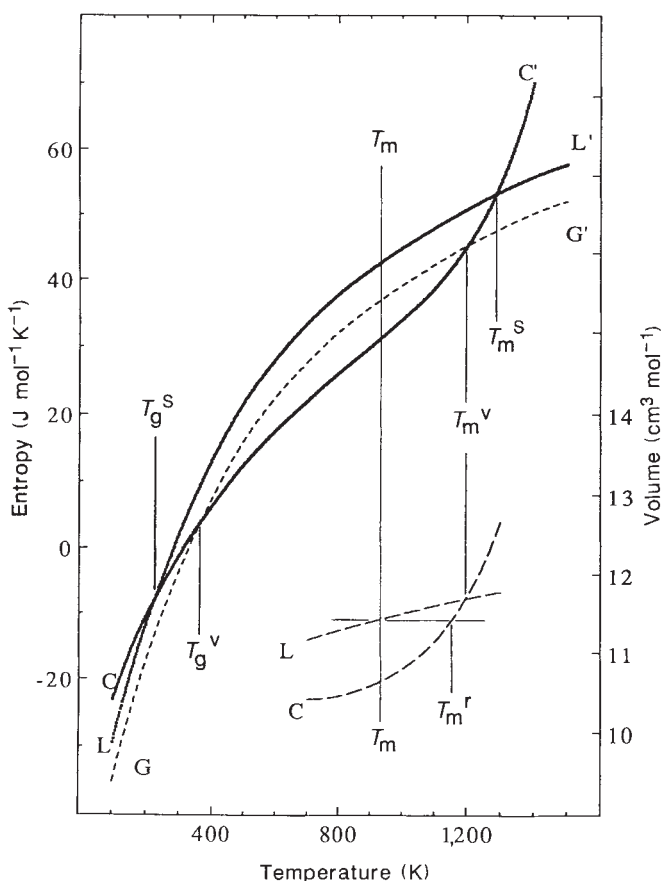


FIG. 2 The calculated entropy and molar volume for aluminium in the crystalline (C) and liquid (L) phases at 1 bar. See text for explanation.

is plotted in Fig. 1 as a function of density. Here a is the lattice parameter and k is Boltzmann's constant. In Fig. 1 the equilibrium thermodynamic melting density and freezing density are marked, together with the instability density at which the superheated crystal spontaneously melts by the generation of dislocations. This instability point corresponds to a value of K close to the value 16π predicted theoretically¹¹. The value of K for the crystal extrapolates to zero at the freezing point density. This means that Born's idea, that melting coincides with the vanishing of a shear modulus, is essentially correct except that this occurs near the freezing point rather than the melting point (note that these do not coincide in T, P, V phase space). We have argued⁸ that the appearance of the communal entropy, once rigidity is lost, allows the system to discontinuously transform from a higher density to this critical density at the freezing point where the rigidity modulus vanishes. The shear instability, properly interpreted, thus naturally gives rise to a first-order melting transition for two dimensions as well as for three dimensions. In conventional isothermal melting the instability occurs within the solid-liquid interface, a region that is effectively superheated as its density is lower than that of the crystal.

Second, we have shown that the entropy change ΔS on isochoric (that is constant volume) melting for a large range of simple crystalline materials is a constant equal to $\sigma R \ln 2$ where σ takes the value 1 for monatomic systems, 2 for binary systems and so on^{12,13}. This has been demonstrated for the inert gases, alkali metals, alkali and alkali earth halides as well as a variety of computer-simulated systems that is, for all systems with approximately spherical atoms. Subsequent investigations of molecular-dynamics-simulated systems showed that this entropy change was just the communal entropy of the dense liquid¹³. Thus, to a good approximation, the only difference in entropy between liquid and crystal at the same density is the communal entropy of the liquid. This reduces to zero in at least three distinct transitions^{13,14}: the glass transition¹⁴ (and as a consequence there is essentially no entropy difference between glass and crystal at the same density), in the superfluid transition¹³ for ⁴He, and in the fast-ion transition of an ionic crystal¹³.

Where liquid and crystal are not at the same density the entropy difference may be well approximated by noting that the isothermal volume dependence of the entropy is given by $(\partial S / \partial V)_T = \alpha_v \beta_T$ where α_v is the volume coefficient of thermal expansion and β_T is the isothermal bulk modulus. The product $\alpha_v \beta_T$ is nearly constant from above the Debye temperature to the melting point and takes nearly the same value in the melt. It is not unexpected, therefore, that we find

$$\Delta S_m = \alpha_v \beta_T \Delta V_m + \sigma R \ln 2 \quad (2)$$

an excellent approximation for the entropy change on melting. This relation is well satisfied for all of the simple systems noted above with approximately spherical atoms¹².

Some simple conclusions using the above results. Figure 2 is a reproduction of Fecht and Johnston's³ diagram of the temperature dependence of the entropy of crystalline (curve C-C') and liquid (curve L-L') aluminium. (Their calculated extrapolations are estimates only and aluminium, because of its non-spherical atoms, is not a good choice for applying the above simple phenomenology. The diagram is, however, appropriate for illustrative purposes.) Fecht and Johnston identified two catastrophe points, marked T_g^s and T_m^s in the figure, at which the entropies of the two phases are equal. By subtracting the communal entropy, $0.69 R$, from the entropy in the liquid phase I obtain the light solid line annotated G-G'. This represents the entropy of the liquid over the entire temperature range if the liquid were non-diffusive, that is, if the atoms were constrained to remain within their individual shells of neighbouring atoms. Such a system is equivalent to a glass. The entropy on this curve, according to equation (2), differs from the crystalline entropy by a volume-dependent term only, and consequently at the two

points marked T_g^v and T_m^v , where the curve G-G' intersects C-C', the crystalline and liquid phases have the same molar volume or density. Beyond these points the crystal would be less dense than its liquid phase at the same temperature and for a close-packed crystal this is quite unexpected. These two points therefore represent further catastrophe points, the boundaries of regions of paradox comparable to those of Kauzmann. (There are materials for which the density of the liquid exceeds that of the crystalline solid⁶, including Ga, Bi, Sb, H₂O and KI at elevated pressure, but these cases arise from the occurrence of a very open crystalline structure.)

Having identified the two isochoric catastrophe points at which crystal and liquid possess the same density and which preempt the isentropic catastrophe points, I can estimate from these entropy curves the molar volumes of the two phases as they vary with temperature. The object of this last calculation is to locate the point of rigidity instability which, according to the above ideas, occurs where the density of the superheated crystal equals that of the liquid at the freezing point. For a non-isothermal process

$$\Delta S = \alpha_v \beta_T \Delta V + \int (C_v/T) dT$$

$$\approx \alpha_v \beta_T \Delta V + C_v \ln(T/T_0) \quad (3)$$

The latter approximation applies if C_v is constant, as may be assumed well above the Debye temperature. Taking T_0 as T_m , the crystalline and liquid volumes are estimated from curves G-G' and C-C' by subtracting $C_v \ln(T/T_m)$ and dividing by $\alpha_v \beta_T$. These are plotted as the dashed lines in Fig. 2, to which

the volume scale on the right-hand side applies. The point T_m^r at which the volume of the superheated crystal equals the volume of the liquid at the freezing point is marked. This is the point at which the rigidity modulus vanishes and shear instability occurs, and this precedes both other catastrophe points T_m^v and T_m^s . I note that the calculations of Boyer¹⁵ indicate that the occurrence of a shear instability drives a preemptive instability in the compressibility. Thus, for crystalline superheating a succession of catastrophes occur in the sequence thermodynamic (equilibrium melting), elastic compressibility, elastic rigidity, isochoric (equal volumes for crystal and liquid) and finally, entropic points. I believe that the last two catastrophes are not observable as they are preceded by the rigidity catastrophe which is the underlying cause of melting. □

Received 2 June; accepted 17 October 1989.

1. Lindemann, F. A. *Z. Phys.* **11**, 609-615 (1910).
2. Born, M. *J. Chem. Phys.* **7**, 591-601 (1939).
3. Fecht, H. J. & Johnson, W. L. *Nature* **334**, 50-51 (1988).
4. Kauzmann, W. *Chem. Rev.* **43**, 219-256 (1948).
5. Cohen, M. H. & Grest, G. S. *Phys. Rev.* **B20**, 1077-1098 (1979).
6. Ubbelohde, A. R. *The Molten State of Matter* (Wiley, Chichester, 1978).
7. Tallon, J. L., Robinson, W. H. & Smedley, S. I. *J. Phys. Chem.* **82**, 1277-87 (1978).
8. Tallon, J. L., Robinson, W. H. & Smedley, S. I. *Nature* **266**, 337-338 (1977).
9. Tallon, J. L. *Phil. Mag.* **A39**, 151-161 (1979).
10. Abraham, F. F. *Phys. Rep.* **80**, 339-374 (1981).
11. Kosterlitz, J. M. & Thouless, D. J. *J. Phys. C*, **5**, L124-L126 (1972); **C6**, 1181-1191 (1973).
12. Tallon, J. L. & Robinson, W. H. *Phys. Lett.* **A87**, 365-368 (1982).
13. Tallon, J. L. *Phys. Rev.* **B29**, 4153-4155 (1984).
14. Madsen, J. U. & Cotterill, R. M. *J. Phys. Lett.* **A83**, 219-220 (1981).
15. Boyer, L. L. *Phase Transitions* **5**, 1-48 (1985).

ACKNOWLEDGEMENTS. I thank the New Zealand DSIR for a study award to visit Cambridge University, where this work was carried out.

Interhemispheric asymmetry in climate response to a gradual increase of atmospheric CO₂

R. J. Stouffer, S. Manabe & K. Bryan

Geophysical Fluid Dynamics Laboratory/NOAA, Princeton University, Princeton, New Jersey 08542, USA

THE transient response of a coupled ocean-atmosphere model to an increase of atmospheric carbon dioxide has been the subject of several studies¹⁻⁸. The models used in these studies explicitly incorporate the effect of heat transport by ocean currents and are different from the model used by Hansen *et al.*⁹. Here we evaluate the climatic influence of increasing atmospheric carbon dioxide using a coupled model recently developed at the NOAA Geophysical Fluid Dynamics Laboratory. The model response exhibits a marked and unexpected interhemispheric asymmetry. In the circumpolar ocean of the Southern Hemisphere, a region of deep vertical mixing, the increase of surface air temperature is very slow. In the Northern Hemisphere of the model, the warming of surface air is faster and increases with latitude, with the exception of the northern North Atlantic, where it is relatively slow because of the weakening of the thermohaline circulation.

The model we use here consists of general circulation models (GCMs) of the atmosphere and oceans and a simple model of the continental surface that involves the budgets of heat and water. It is a global model and includes realistic geography. The atmospheric component of the model used here is very similar to the model used in ref. 10. The seasonally varying insolation at the top of the model atmosphere is taken into account in calculating the solar radiation. Whenever the relative humidity at a grid point exceeds a critical value (99%), the grid box is filled with cloud; otherwise, it is free of cloud. The atmospheric GCM uses a spectral transform method in which the horizontal distributions of predicted variables are represented by spherical

harmonics (15 associated Legendre functions for each of 15 Fourier components) and grid-point values. For vertical-finite-difference calculations, nine unequally spaced levels are chosen. The ocean GCM uses a full finite-difference technique and has a regular grid system with 4.5° × 3.75° (latitude × longitude) spacing and 12 vertical levels¹¹. The atmospheric and oceanic components of the model interact with each other continuously through the exchange of heat, water and momentum fluxes.

Two 100-year integrations of the coupled model were performed. In the first integration the atmospheric concentration of carbon dioxide increased by the rate of 1% yr⁻¹ (compound). The CO₂ concentration remains unchanged in the second integration (the control). We have evaluated the influence of

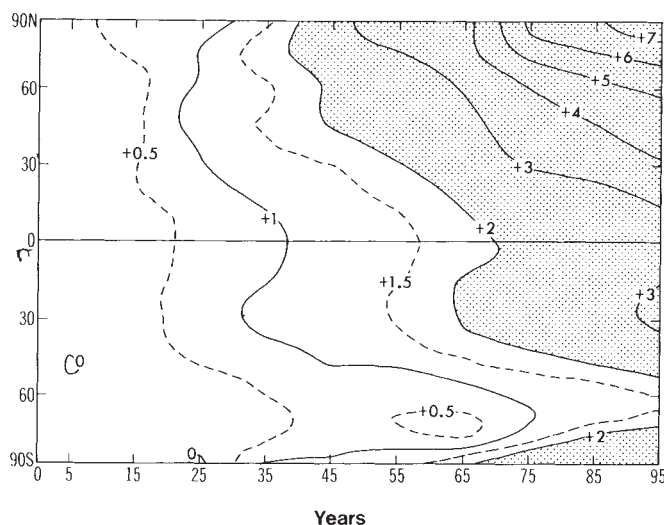


FIG. 1 Latitude-time distribution of the difference in zonally averaged, decadal-mean surface air temperature (°C) between the two integrations. Zonal-mean surface air temperature is averaged over the decades centred at the fifth, fifteenth . . . , and ninety-fifth year of the experiment.

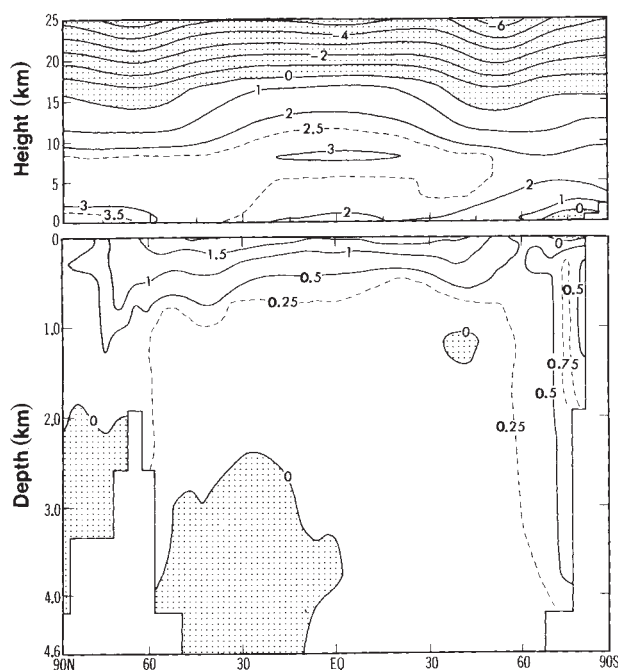


FIG. 2 Zonal-mean difference in temperature ($^{\circ}\text{C}$) of the ocean-atmosphere model between the two integrations. The difference represents the decadal average over years 61–70. The atmospheric data have been interpolated to isobaric surfaces.

gradually increasing atmospheric carbon dioxide upon the coupled system from the difference between the first and second (control) integrations. If the CO_2 increase of $1\% \text{ yr}^{-1}$ begins in AD 1958, the doubling of CO_2 concentration would be achieved around AD 2030. Such an increase of CO_2 is larger than its projected increase but is approximately equal to the present rate of increase of the combined thermal forcing of all greenhouse gases (other than water vapour)¹².

The initial conditions for both integrations have realistic seasonal and geographical distributions of surface temperature, surface salinity and sea ice, with which both the atmospheric and oceanic components of the model are nearly in equilibrium. This quasi-equilibrium condition was obtained by separate time integrations of these two components of the model using observed surface conditions. The convergence of the oceanic component towards equilibrium was accelerated by the method described in ref. 13.

To avoid the systematic climate drift caused by the bias of the model during the two 100-year integrations, the oceanic surface fluxes of heat and water are adjusted seasonally and

geographically, but not interannually, using the method described in ref. 14. Identical adjustments are applied to both the first and the second integrations defined above. By contrast to recent studies^{4,8}, no systematic trends are evident during the control integration, indicating that the initial condition is close to stable equilibrium and the flux adjustments are working as expected.

Figure 1 shows the latitude-time distribution of the CO_2 -induced change of zonally averaged, decadal-mean surface air temperature during the 100-year experiment. In the Northern Hemisphere the warming of surface air increases with latitude, partly because of the poleward retreat of both snow cover and sea ice which have high surface albedo. The relatively large warming in high northern latitudes is essentially confined to the lower troposphere because of the stable stratification. This is indicated in Fig. 2 which shows the difference in zonal-mean temperature between the two integrations for the seventh decade of the experiment. The ocean warming is essentially confined to the upper layer, except near Antarctica. It also increases with latitude in the Northern Hemisphere, but becomes small near the North Pole which is covered by sea ice.

The meridional profile of the near-surface warming in the Northern Hemisphere is qualitatively similar to the difference between two equilibrium climates of a model with the normal and the above-normal concentrations of atmospheric CO_2 . (See, for example, refs 15, 16, studies that were conducted using an atmosphere/mixed-layer-ocean model and atmosphere-ocean GCM, respectively.)

Over the northern North Atlantic, the increase of surface air temperature is significantly smaller than the zonal average. This is evident in Fig. 3 which shows the geographical distribution of the difference in decadal-mean surface air temperature between the two integrations for the seventh decade of the experiment. Because of the increase in runoff and precipitation over evaporation, surface salinity decreases and the stability of upper ocean layer increases in the northern North Atlantic. Thus, the thermohaline circulation of the North Atlantic is weakened by $\sim 25\%$ by the seventh decade of the experiment. The weakening is shown in Fig. 4 which illustrates the meridional circulation zonally averaged over the entire oceanic segment of a latitude circle. (Note the reduction from 16 to 12 Sv in the Northern Hemisphere.) Because of the weakening of the thermohaline circulation, the northward advection of warm and saline subtropical surface water is reduced, thereby reducing surface salinity further and making the increase of sea surface temperature relatively small over the northern North Atlantic Ocean as indicated in Fig. 3. The reduction of surface salinity stabilizes the near-surface layer and provides a feedback that further weakens the thermohaline circulation in the northern North Atlantic.

In an earlier study, Bryan and Spelman³ investigated the

FIG. 3 The geographical distribution of the difference in surface air temperature ($^{\circ}\text{C}$) between the two integrations averaged over years 61–70.



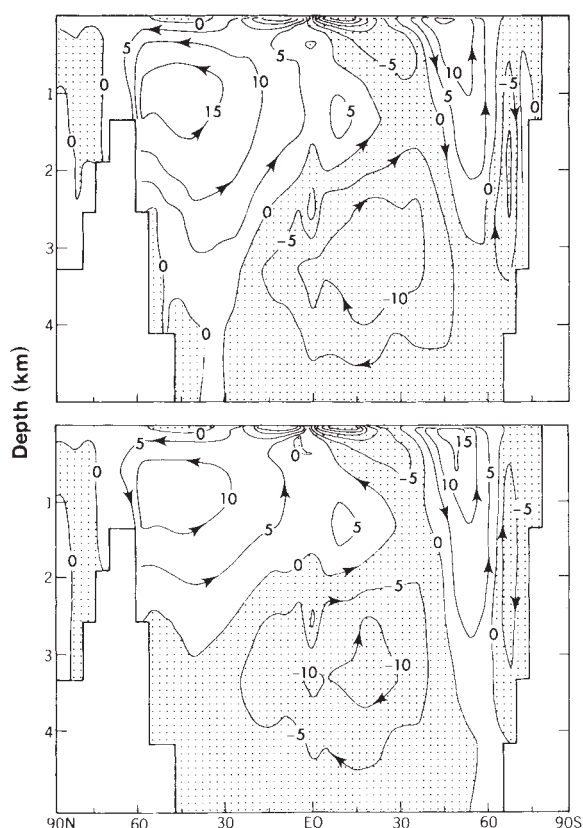


FIG. 4 Streamlines of zonal-mean meridional oceanic circulation (in Sverdrups) averaged over the seventh decade (that is, years 61–70) of the experiment. Top, the control integration with constant CO_2 ; bottom, the integration in which atmospheric CO_2 concentration is increased with the rate of $1\% \text{ yr}^{-1}$.

influence of an abrupt quadrupling of atmospheric CO_2 on the climate of an atmosphere–ocean model with a sector computational domain bounded by the Equator and two meridional boundaries. They found that the thermohaline circulation of the model collapsed completely towards the end of the third decade of the experiment. Washington and Meehl⁸ noted that the response of their global model to a linear increase of atmospheric CO_2 was a weakening of the thermohaline circulation in the Atlantic. Manabe and Stouffer¹⁴ found that their global model has two stable equilibria, one with and one without the thermohaline circulation in the Atlantic Ocean. Although we do not observe a complete collapse of the thermohaline circulation here, it is a possibility that deserves¹⁷ further assessment.

In the Southern Hemisphere, the CO_2 -induced warming of surface air decreases with increasing latitude in sharp contrast to the situation in the Northern Hemisphere and to the results of equilibrium experiments. The rate of the warming is particularly slow in the Antarctic circumpolar ocean. A similar feature is noted and analysed in studies^{6,7} that explored the

effect of an abrupt doubling of the atmospheric CO_2 on coupled models with 'sector' and global computational domain, respectively. The latter model is very similar to the present model except that it has annual-mean insolation. The fraction of the area covered by oceans in the Southern Hemisphere is much larger than the corresponding fraction in the Northern Hemisphere. Thus the effective thermal inertia of the Southern Hemisphere tends to be larger than in the Northern Hemisphere¹⁸, accounting for the slower rise of surface air temperature in the former.

Based upon a detailed study of the heat budget^{6,7} there is another important mechanism responsible for the interhemispheric difference of effective thermal inertia particularly in high latitudes. From 45° – 65° S, the strong surface westerlies induce an Ekman drift current towards the Equator. Because of the absence of a continental barrier in the upper oceanic layer, the zonal pressure gradient is small at the latitudes of the Drake Passage and the southward geostrophic return flow is weak beneath the surface mixed layer. Instead, the surface drift currents are compensated by a deep downwelling north of the Drake Passage gap, southward geostrophic flow at the base of the gap, and deep upwelling south of the circumpolar current, as indicated in Fig. 4. The existence of such an upwelling region is inferred^{19,20} from water mass analysis. A deep wind-driven cell is also simulated by eddy-resolving, high-resolution models²¹. In addition to the wind-driven deep cell of meridional circulation described above, we note that (Fig. 4) there is another cell circulating in the reverse direction in the immediate vicinity of Antarctica where deep convective overturning prevails, forming the Antarctic bottom water. These deep cells of meridional circulation together with associated convective overturning and other sub-grid-scale mixing processes enhance the vertical mixing of heat over a thick ocean layer, thereby increasing the effective thermal inertia of the coupled ocean–atmosphere system. This large effective thermal inertia is responsible for the very slow rise of the near-surface temperature in the circumpolar ocean of the model. It is likely that a similar effective vertical mixing is also realized in the actual circumpolar ocean which is characterized by relatively small vertical variations of density and other trace substances such as phosphate and ^{14}C .

The marked interhemispheric asymmetry of our result differs from the study of Schlesinger *et al.*⁴ in which an almost symmetric response was obtained. A recent result⁸ indicates smaller warming in the Southern than in the Northern Hemisphere. However, the asymmetry is much less pronounced than in the present results. This discrepancy may result from the fact that the wind-driven cell in the circumpolar ocean obtained in the earlier studies is much weaker and shallower than the cell simulated by the present model. Obviously, a careful comparison of results is needed to determine the underlying cause of this discrepancy.

Further evaluation of the mechanisms involved in the effect of ocean circulation on the CO_2 -induced climate change is in progress. The general applicability of the present results will be evaluated by conducting further simulations with different rates of CO_2 change and with models of higher computational resolution. □

Received 29 June; accepted 19 October 1989.

1. Bryan, K., Komro, F. G., Manabe, S. & Spelman, M. J. *Science* **215**, 56–58 (1982).
2. Spelman, M. J. & Manabe, S. *J. geophys. Res.* **89**, 571–586 (1984).
3. Bryan, K. & Spelman, M. J. *J. geophys. Res.* **90**, 11679–11688 (1985).
4. Schlesinger, M. E. & Jiang, X. *Climate Dynamics* **3**, 1–17 (1988).
5. Schlesinger, M. E., Gates, W. L. & Han, Y. J. in *Coupled Ocean–Atmosphere Models* (ed. Nihoul, J. C. J.) 447–478 (Elsevier, Amsterdam, 1985).
6. Bryan, K., Manabe, S. & Spelman, M. J. *J. phys. Oceanogr.* **18**, 851–867 (1988).
7. Manabe, S., Bryan, K. & Spelman, M. J. *J. phys. Oceanogr.* (in the press).
8. Washington, W. & Meehl, G. *Clim. Dynamics* **2**, 1–38 (1989).
9. Hansen, J. *et al.* *J. geophys. Res.* **93**, 9341–9364 (1988).
10. Wetherald, R. T. & Manabe, S. *J. Atmos. Sci.* **45**, 1397–1415 (1988).
11. Bryan, K. & Lewis, L. J. *J. geophys. Res.* **84**, 2503–2517 (1979).

12. Ramanathan, V. R., Cicerone, R. J., Singh, H. B. & Kiehl, J. T. *J. geophys. Res.* **90**, 5547–5566 (1985).
13. Bryan, K. *J. phys. Oceanogr.* **14**, 666–673 (1984).
14. Manabe, S. & Stouffer, R. J. *J. Clim.* **1**, 841–866 (1988).
15. Manabe, S. & Stouffer, R. J. *J. geophys. Res.* **85**, 5529–5524 (1980).
16. Manabe, S. & Bryan, K. *J. Geophys. Res.* **90**, 11689–11707 (1985).
17. Broecker, W. S. *Nature* **329**, 123–125 (1987).
18. Thompson, S. L. & Schneider, S. H. *Science* **217**, 1031–1033 (1982).
19. Deacon, G. E. R. *Discovery Rep.* **15**, 125–152 (1937).
20. Sverdrup, H. V., Johnson, M. W. & Fleming, R. H. *The Oceans* (Prentice Hall, Englewood Cliffs, 1942).
21. Semtner, A. J. & Chervin, R. M. *J. geophys. Res.* **93**, 15502–15522 (1988).

ACKNOWLEDGMENTS. We thank I. Held, N.-C. Lau, J. D. Mahlman and R. Toggweiler for their comments on the manuscript and J. D. Mahlman for his support.

Cause of sulphate attack on concrete, render and stone indicated by sulphur isotope ratios

K. Pye* & N. Schiavon†

* Postgraduate Research Institute for Sedimentology, University of Reading, Whiteknights, Reading RG6 2AB, UK

† Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK

SULPHATE attack is a significant problem which locally affects concrete and other construction materials in the United Kingdom and overseas. The mechanisms of sulphate attack are fairly well understood, but in some instances it has proved difficult to establish the origin of the sulphate. Possible sulphate sources include soils and ground water, sea water, atmospheric pollution, contaminated ground-fill and highway de-icing salts. Here we show that stable-sulphur-isotope analysis can help to identify the causes of sulphate attack. We carried out this investigation primarily to determine whether sulphate impurities in rock salt used for winter maintenance of the roads could be responsible for higher than expected sulphate levels that have been observed in parts of some road bridges. Our sulphur isotope results indicate that this is not the case, and indicate that atmospheric sulphate sources are more important.

Sulphate attack can cause serious deterioration of hardened concrete, render and mortar¹⁻³. Both chemical and mechanical mechanisms may be involved. In the presence of dissolved sulphates, the constituents of hardened Ordinary Portland Cement (OPC) paste are progressively chemically altered to form secondary minerals such as ettringite, thaumasite and gypsum^{3,4}. These changes can cause expansion, cracking and loss of cement-like characteristics and may ultimately lead to total failure of the structure. The extent and rate of chemical alteration is dependent on the concentration and type of sulphate in solution. Sulphate concentrations as low as 0.3 g l⁻¹ can pose a significant threat to concrete made with OPC⁵. Magnesium and ammonium sulphate solutions pose the greatest threat because they react chemically with all of the compounds present in hydrated OPC. Sodium sulphate solutions only attack the tricalcium aluminate hydrate and calcium hydroxide components, whereas calcium sulphate solutions react only with tricalcium aluminate hydrate^{1,2}.

In some circumstances three physical mechanisms also contribute to the deterioration⁶⁻⁸: (1) stress generated by salt crystallization from a saturated or supersaturated solution; (2) stress generated by thermal expansion of salt crystals; (3) stress caused when certain salts undergo an increase in volume during hydration. The relative importance of these mechanisms is dependent on the nature of salts present, on fluctuations in ambient temperature and humidity, and on the physical characteristics of the material being attacked^{7,8}.

In assessing the likelihood or cause of sulphate attack it is therefore important to identify the type, concentration and origin of sulphates that may come in contact with the structure.

Sulphate can be derived from several sources including naturally saline soils and ground water, sea water, atmospheric pollution, industrial effluents, ash and other solid wastes used as foundation material, oxidation of pyritic shales, oxidation of sulphide inclusions in concrete aggregates, highway de-icing salts, agricultural fertilizers and effluents, and bacterial degradation of sewage sludge⁹⁻¹². Chemical analysis of soils and ground waters, combined with petrographic examination of the concrete, normally allows some possible sulphate sources to be eliminated. Here we show that stable-sulphur-isotope analysis

can also assist in identifying the source of sulphate that has invaded concrete and related structures.

The near-surface concrete in road bridges is quite frequently found to contain higher than expected levels of sulphate. The cement in hardened OPC concrete normally contains 2.5-2.9% (average 2.7%) sulphate expressed as SO₃, but sulphate levels sometimes exceed 4.5% SO₃ by weight of cement in the outer parts of concrete structures that are 15-20 years old¹³. On parts of the structures that are not in contact with soils or ground water, locally high sulphate levels might be explained by one of three factors: (1) redistribution of sulphate originally present in the concrete by fluid migration; (2) ingress of sulphates present in de-icing salt solutions; (3) ingress of sulphur species from the atmosphere. Here we identify the relative importance of these three possible sources in contributing to high sulphate levels that were found during an investigation of a number of Cambridgeshire bridges¹³.

Most of the salt used for winter maintenance of the roads in the United Kingdom, including Cambridgeshire, is Triassic rock salt from the ICI Meadowbank mine at Winsford, Cheshire¹⁴. This salt contains sulphate impurities mainly in the form of anhydrite (CaSO₄). Anhydrite is also the dominant sulphate impurity in de-icing salt from other sources used in the United Kingdom.

To determine whether rock salt has contributed significantly to higher than expected sulphate levels found in the Cambridgeshire bridges studied, $\delta^{34}\text{S}$ values (defined in Table 1 legend) were determined for 44 samples of concrete, de-icing salt and other sulphate-bearing materials found in association with degraded concrete from a number of known environments. A full list of the samples is given in Table 1.

All of the samples were oven dried at 80 °C for 24 h, ground to pass a 0.1-mm sieve and boiled in 10% (v/v) HCl for 20 min to extract the sulphate. After filtering, the acid-soluble sulphate was precipitated by addition of 5% (w/v) BaCl₂ solution. The resulting BaSO₄ precipitates were then washed with de-ionized water and dried at 80 °C. The $^{34}\text{S}/^{32}\text{S}$ ratios were determined by mass spectrometry performed on the SO₂ obtained by heating the BaSO₄ to ~1,400 °C (refs 15, 16). The overall reproducibility of a BaSO₄ standard was better than $\pm 0.3\%$.

The results showed the following (Fig. 1): (1) samples of unhydrated OPC and pristine hardened concrete obtained from cores in Cambridgeshire bridges have $\delta^{34}\text{S}$ values of +10 to +14.5‰; (2) the samples of black ash, shale, clay and brick, which contain reduced sulphur compounds or oxidation prod-

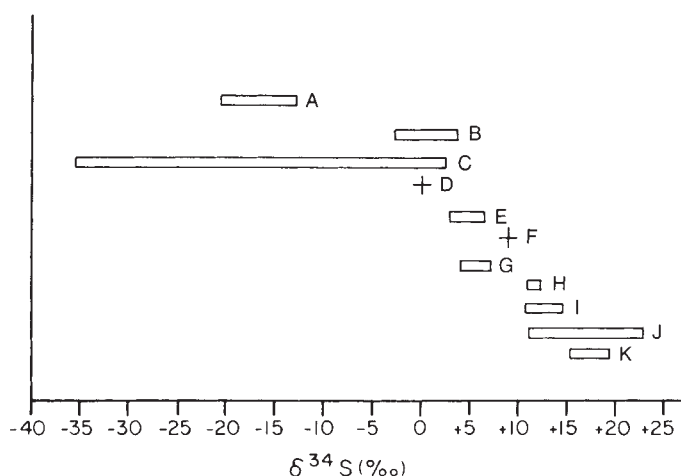


FIG. 1 $\delta^{34}\text{S}$ values for: A, brick (flettons); B, render affected by sulphate from brick/atmospheric sources; C, clay, shale and industrial ash ground-fill; D, concrete affected by sulphate from ash-fill; E, gypsum crusts on limestone buildings; F, sewer concrete affected by sulphuric trade effluent; G, deteriorated concrete, affected by atmospheric sulphate; H, pristine concrete, Peterborough; I, unhydrated OPC; J, other UK and overseas de-icing salt; K, Cheshire de-icing salt.

TABLE 1 Results of sulphur isotope analysis

Sample No.	Nature of material	$\delta^{34}\text{S} (\text{‰})^*$	Sample No.	Nature of material	$\delta^{34}\text{S} (\text{‰})^*$
Unhydrated cement			Stone weathering crusts		
KP3	Blue Circle OPC	+14.5	KP23	gypsum crust on magnesian limestone, Westminster	+6.8
KP4	Rugby OPC	+10.6	KP1	gypsum crust on limestone, King's College, Cambridge	+4.0
Pristine hardened concrete from bridge cores			KP2	gypsum crust on limestone, Pembroke College, Cambridge	+3.5
KP13	depth 18 cm, Peterborough (Bretton)	+11.9	Deteriorated concrete and render		
KP14	depth 18 cm, Peterborough South	+11.1	KP10	deteriorated concrete affected by sulphates, Bretton Interchange, depth 0–5 cm, Peterborough	+5.7
KP12	depth 15 cm, Peterborough (Fulbridge)	+11.2	KP25	as above	+7.1
Cheshire rock salt			KP11	deteriorated concrete affected by sulphates, Fulbridge footbridge deck, Peterborough, depth 0–5 cm	+6.9
WIN1		+15.2	KP27	calcite stalactite below soffit, Fulbridge footbridge	+4.7
WIN2		+17.5	KP24	carbonated concrete pillar protected from rainfall, underground car park, Huntingdon	+10.2
WIN3		+17.5	KP5	concrete paving slabs exposed to atmospheric sulphate but not treated with de-icing salt, Wolverhampton	+7.0
WIN4		+18.7	KP6	as above	+7.4
WIN5		+18.7	KP7	concrete cladding panels affected by atmospheric sulphate, 6th floor flat, Leeds	+5.9
WIN6		+18.9	KP8	as above	+6.9
De-icing salt from other UK and overseas sources used in UK			KP26	sewer concrete attacked by sulphates and sulphuric acid in trade effluent, Cambridge	+8.7
C1	sources confidential	+19.3	KP28	concrete floor slab attacked by sulphate from black ash fill, Birmingham	0.0
C2		+22.8	KP9	render attacked by sulphates from brick, Cambridge University Computer building	–2.7
C3		+17.9	KP43	render attacked by sulphates, partly brick and partly atmospheric derived, Fowlmere	+3.8
C4		+11.5			
C5		+10.2			
C6		+11.9			
C7		+11.3			
Brick					
KP33	Fletton KN	–20.4			
KP34	Fletton BN	–14.9			
KP35	Fletton CN	–13.3			
Clay, shale and ash ground fill					
KP36	Oxford Clay, King's Dyke	–10.7			
KP37	Oxford Clay, Calvert	+1.2			
KP38	Oxford Clay, Bletchley	–1.8			
KP31	oxidized Edale Shale, Carsington	–31.6			
KP32	oxidized Edale Shale, Carsington	–36.0			
KP29	black industrial ash, Birmingham	+2.6			
KP30	red industrial ash, Birmingham	–0.4			

* $\delta^{34}\text{S} = [({}^{34}\text{S}/{}^{32}\text{S})_{\text{sample}}/({}^{34}\text{S}/{}^{32}\text{S})_{\text{standard}}] - 1$, where we have used the Canyon Diablo Troilite (CDT) standard.

ucts derived from them, have $\delta^{34}\text{S}$ values in the range –36 to +2.6‰; (3) concrete samples known to have suffered contamination by ingress of sulphates derived from the above sources have $\delta^{34}\text{S}$ values of –7 to +3.8‰; (4) gypsum crusts that have formed on the upper parts of limestone buildings in Cambridge because of the deposition of atmospheric sulphate have $\delta^{34}\text{S}$ values of +4 to +6.8‰; (5) deteriorated concrete from high levels on blocks of flats, which have also been affected by ingress of atmospheric sulphate, have $\delta^{34}\text{S}$ values of +5.9 to +7.0‰; (6) Cheshire rock salt supplied for winter maintenance of the roads has typical $\delta^{34}\text{S}$ values of +15.2 to +18.9‰; (7) samples of de-icing salt from a number of other mines in the United Kingdom and overseas, which are used less widely for winter maintenance, possess $\delta^{34}\text{S}$ values of +10.2 to +22.8‰; (8) samples of deteriorated concrete from bridges in the Peterborough area, which were found to have higher than expected sulphate levels but where the source of the extra sulphate was uncertain¹³, has $\delta^{34}\text{S}$ values of +4.7 to +7.1‰.

These results clearly indicate that de-icing salt could not have contributed significantly to the elevated sulphate levels in the Cambridgeshire bridges investigated, because the rock salt has more positive $\delta^{34}\text{S}$ values, and the sulphate-contaminated concrete lower positive values, than the pristine hardened concrete, which has not been affected by ingress of external sulphate, at depths of 150–180 mm in these structures. The results also indicate that the high sulphate levels in the outer 40 mm of concrete cannot be attributed to redistribution, during wetting and drying, of sulphate initially present within the concrete itself. As the near-surface concrete with high sulphate levels has lower $\delta^{34}\text{S}$

values than the pristine concrete at greater depth, the extra sulphate is most likely to have been introduced from atmospheric sources. This may include both dry deposition of sulphur dioxide and wet deposition of sulphur species in rainfall and spray thrown up by vehicles on the carriageway.

We conclude that sulphur isotope analysis can provide a useful tool in investigations to determine the causes of sulphate attack. Further work is required to determine whether atmospheric sulphur deposited by wet and dry processes can be differentiated and whether the incorporation of external sulphate in stone crusts and concrete involves any significant differences in isotope fractionation. □

Received 22 September; accepted 6 November 1989.

1. Figg, J. *Chem. Ind.* 17 October, 770–775 (1983).
2. Benstead, J. *World Cement Technol.* May, 178–184 (1981).
3. Mehta, P. K. *Cement Concrete Res.* **13**, 401–406 (1983).
4. Crammond, N. J. *Cement Concrete Res.* **15**, 431–441 (1985).
5. BRE Digest **250** *Concrete in Sulphate-bearing Soils and Groundwaters*. (Building Research Establishment, Garston, 1981).
6. Cooke, R. U. & Smalley, I. J. *Nature* **220**, 1226–1227 (1968).
7. Sperling, C. H. B. & Cooke, R. U. *Earth Surface Processes and Landforms* **10**, 541–555 (1985).
8. Goudie, A. S. *Sch. Geogr. Res. Pap.* **33**, (University of Oxford, 1986).
9. Lossing, F. A. *Highway Res. Board Rec.* **113**, 88–103 (1966).
10. Mour, J. & Rosenquist, I. T. J. *Am. Concrete Inst.* September, 257–264 (1959).
11. Cheetham, D. W. & Banfill, P. F. G. *Construction Pap.* **1**, 7–16 (1980).
12. St John, D. A. *Cement Concrete Res.* **12**, 633–639 (1982).
13. Pye, K. & Hudson, A. J. *New Civil Engrn* 10 July, 16–17 (1986).
14. Mineral Resources Consultative Committee *Salt* (HMSO, London, 1971).
15. Holt, B. D. & Engelkeimer, A. G. *Analyt. Chem.* **42**, 1451–53 (1970).
16. Faure, H. *Principles of Isotope Geochemistry* 2nd edn, 523–552 (Wiley, New York, 1986).

ACKNOWLEDGEMENTS. We thank A. J. Hudson, A. E. Fallick and A. Boyce for their help. University of Reading PRIS Contribution No. 043.

Thermal control of basaltic fissure eruptions

Paul M. Bruce* & Herbert E. Huppert

Institute of Theoretical Geophysics, University of Cambridge,
20 Silver Street, Cambridge CB3 9EW, UK

* Present address: Shell International Petroleum Co. Ltd., Shell Centre,
London SE1 7NA, UK

AS hot, basaltic magma rises in a newly opened dyke during an eruption, it flows through colder crustal rock. The flowing magma advects heat along the dyke; at the same time, heat is conducted out of the dyke into the colder surroundings. The loss of heat will lead initially to the channel becoming constricted by magma solidifying against the walls. The channel may then become completely blocked, which will end the eruption at that site before the supply of magma is exhausted. Alternatively, the continual supply of heat by the flowing magma may, after some time, exceed the losses into the country rock. In this case, the initial solidification is reversed, the walls of the channel are progressively melted and the dyke is widened until the supply diminishes. We present here a model that quantitatively delineates these two regimes. We also identify an intermediate regime in which parts of the surface fissure may become blocked and the eruption continues from isolated vents.

Basaltic eruptions display a fascinating variety of behaviours¹. The low viscosity of such lavas allows relatively steady degassing and flow, as opposed to the explosive eruptions characteristic of more viscous magmas. 'Quiet' basaltic eruptions nonetheless evolve over time in a highly complex way. Typically, a linear system of fissures opens rapidly and erupts a continuous 'curtain of fire'. Small eruptions simply cease within half a day. In larger eruptions activity begins, after several hours, to concentrate at certain points along the fissure^{2,3}. If eruption persists, it becomes localized at one or a few vents, building cones. Preferential localization of subsequent fissure eruptions to the same areas produces substantial shield volcanoes. In the extremely voluminous eruption at Laki in 1783, flow persisted at not one but dozens of vents along the fissure⁴.

The initial formation of the dyke is the result of elastic processes in the surrounding country rock^{5,6}. We concentrate

To obtain our results we analysed an initial-value problem for $\theta(x, z, t)$, the temperature in both the fluid and the solid with respect to locally cartesian coordinates z measured along the wall and x perpendicular to it, with $x=0$ at the wall. In constructing the equations, we assumed that the flow in the thermal boundary can be represented by a shear flow of magnitude γ . The governing equations then become

$$\theta_t - v\theta_x + \gamma x\theta_z = k\theta_{xx} \quad (x, t > 0) \quad (1)$$

$$\theta = T_m(z=0), \quad \theta = T_w(x=0), \quad \theta = T_m(t=0 \text{ or } x \rightarrow \infty) \quad (2)$$

$$\theta_t - v\theta_x = k\theta_{xx} \quad (x < 0, t > 0) \quad (3)$$

$$\theta = T_w(x=0), \quad \theta = T_\infty(t=0 \text{ or } x \rightarrow -\infty) \quad (4)$$

The velocity of migration of the interface between fluid and solid, v , is proportional to the difference in the conductive heat flux across the wall and is determined from

$$(L/c)v = -\kappa[\theta_x(0+, z, t) - \theta_x(0-, z, t)] \quad (5)$$

The shear flow at the edges of the channel of width W is driven by a pressure difference ΔP and results in a volume flow rate $Q(t)$. These are related by

$$\Delta P = 12\mu Q(t) \int_0^h W^{-3}(z, t) dz \quad \text{and} \quad \gamma(z, t) = 6W^{-2}(z, t)Q(t) \quad (6)$$

With either solidification or melting, the half-width gradually changes according to

$$W(z, t) = W_i - 2 \int_0^t v(z, t') dt' \quad (7)$$

The solutions to equations (1)–(7) were obtained numerically by integrating suitably simplified and non-dimensionalized versions which are formally valid for all time^{10,11}.

here on the primarily thermodynamic problem of determining both the flow in the dyke and its evolution after the initial rock propagation and widening have taken place. Naive estimates of the solidification of magma against the walls of the dyke based solely on conduction⁷, indicate that flow should effectively cease in a matter of days. A previous model⁸ which included some advective effects came to much the same conclusion unless brecciation of the dyke wall and local widening of the conduit by removal of the fractured blocks leads to increased flow rates⁹. There are deficiencies in this model, however, most notably the neglect of both the latent heat of solidification and the effect of the thermal advection in the magma on the temperature profile of the solid. These omissions consistently overestimate the tendency of the dyke to become blocked. Our model includes both these effects and allows us to understand how meltback of the walls can actually occur.

We consider, as depicted in Fig. 1, the dyke to be a two-dimensional channel the width of which is initially uniform but subsequently changes because of local solidification or melting. The newtonian flow is considered to be laminar and the driving pressure is taken to be fixed. Thus if the channel becomes constricted, the total flow rate decreases. Cooling of the magma in the dyke is typically confined to a thin boundary layer adjacent to each wall. The width of the boundary layer increases with length along the dyke and can attain a value of ~ 10 cm at the end of the dyke. The difference between the heat supplied to the wall from this boundary layer and that conducted into the surrounding solid determines the rate of solidification or melting, which is directly proportional to the magnitude of the latent heat. The temperature profile in the solid is determined numerically taking into account the effects of advection in the magma. A full derivation of the governing equations and details of their solution are given in refs 10 and 11.

The behaviour of the solutions is governed by three dimensionless parameters. The first is the Stefan number of the magma, $S_m = L/(c(T_m - T_w))$, where L and c are the latent and specific heats per unit mass, T_m is the temperature of the magma

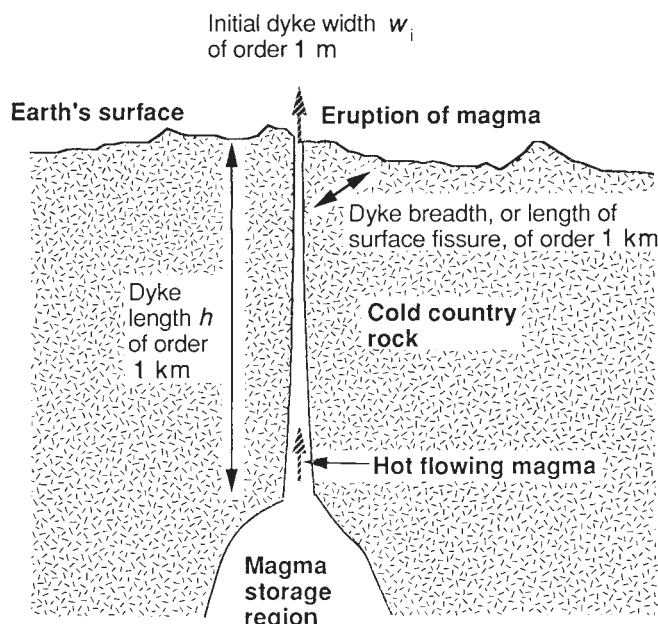


FIG. 1 Cross-section through a typical sheet-like planar dyke. Note that the dyke need not necessarily be vertical.

at the entrance to the dyke and T_w is the temperature at the wall, which is defined as the temperature at which the crystallizing magma ceases to flow. In fact, solidification and melting take place over a temperature range; this range is, however, typically small compared with the difference between T_w and T_∞ , the far-field temperature in the crust, and hence can be neglected. The second parameter governing the behaviour of the thermal processes is the Stefan number of the surrounding crust $S_\infty = L/(c(T_w - T_\infty))$. This incorporates the approximation that the fluid magma, its solid product and the country rock have identical thermal properties. The final parameter is $B = (W_i^4/H^2)(\Delta P/\mu\kappa)$, where W_i is the initial width of the dyke of length H , ΔP is the driving pressure in the reservoir, μ is the mean value of the dynamic viscosity of the magma within the boundary layer and κ is the thermal diffusivity of both magma and country rock. The cube root of B is the ratio of the initial width of the dyke to the initial width, at the top of the dyke, of the thermal boundary layer, due to advection before any solidification or melting can take place. We assume that this ratio is large.

We also assume that $c = 730 \text{ J kg}^{-1} \text{ }^\circ\text{C}^{-1}$, $L = 8 \times 10^5 \text{ J kg}^{-1}$, $\kappa = 10^{-6} \text{ m}^2 \text{ s}^{-1}$, $T_m = 1,200 \text{ }^\circ\text{C}$, $T_w = 1,150 \text{ }^\circ\text{C}$ and $\mu = 100 \text{ Pa s}$. (ref. 12), which reflect values appropriate for a typical basaltic magma. The driving pressure was evaluated from the lithostatic overburden¹³ as $\Delta P/H = 2,000 \text{ N m}^{-3}$, with H taken as either 2 or 5 km. We chose these values because the depth interval 2–5 km corresponds to the top and bottom of active magma reservoirs in Hawaii, Iceland and the East Pacific Rise. It also corresponds approximately to the upper surface and the keel of dykes in the respective rift zones of these localities. In a given volcanic system the remaining two variables T_∞ and W_i determine the course of successive eruptions. The value of T_∞ may vary between 0 and $1,150 \text{ }^\circ\text{C}$ (T_w) depending on the previous

thermal history of the dyking zone. With these values, $S_m = 22$, S_∞ ranges upwards from 1.05 (which corresponds to cold country rock at $0 \text{ }^\circ\text{C}$) and B can take on any positive value, though it must exceed unity for our boundary-layer model to be appropriate.

The important qualitative behaviour of our two-dimensional model is sketched in Fig. 2. This is augmented by the quantitative results depicted in Fig. 3. When the eruption commences magma initially solidifies against the walls of the dyke. In those dykes for which W_i and T_∞ are sufficiently small so that they fall below demarcation line AA' for a 2-km dyke and BB' for a 5-km dyke, solidification continues in the downstream portion of the channel. The surface fissure will eventually become blocked and the eruption through the dyke feeding the fissure will cease irrespective of the magma supply. If $T_\infty = 100 \text{ }^\circ\text{C}$, for a dyke 2 km long the critical initial width is 1.3 m, whereas for a dyke 5 km long the critical initial width is 1.7 m. The time taken for the dyke to become blocked as a function of its initial width is shown in Fig. 3b for these two dyke lengths.

If ΔP remains fixed, eruptions can only continue from dykes whose initial width exceeds the critical value, which is decreased if the dyke is surrounded by warmer country rock. This increase in T_∞ could be the result of pre-warming by previous dyking events or by hydrothermal circulation. For dyke widths beyond the critical value the thermal energy from the reservoir advected along the dyke melts back the walls and the channel remains open. Thereafter, the rate and duration of the eruption are controlled by the decay of the driving pressure ΔP and the possible closing of the dyke because of the resulting subsidence, which we have not included in our model. When ΔP has decreased sufficiently solidification will take place once more and the eruption fade away.

The concept of the two-dimensional widening of a dyke as a

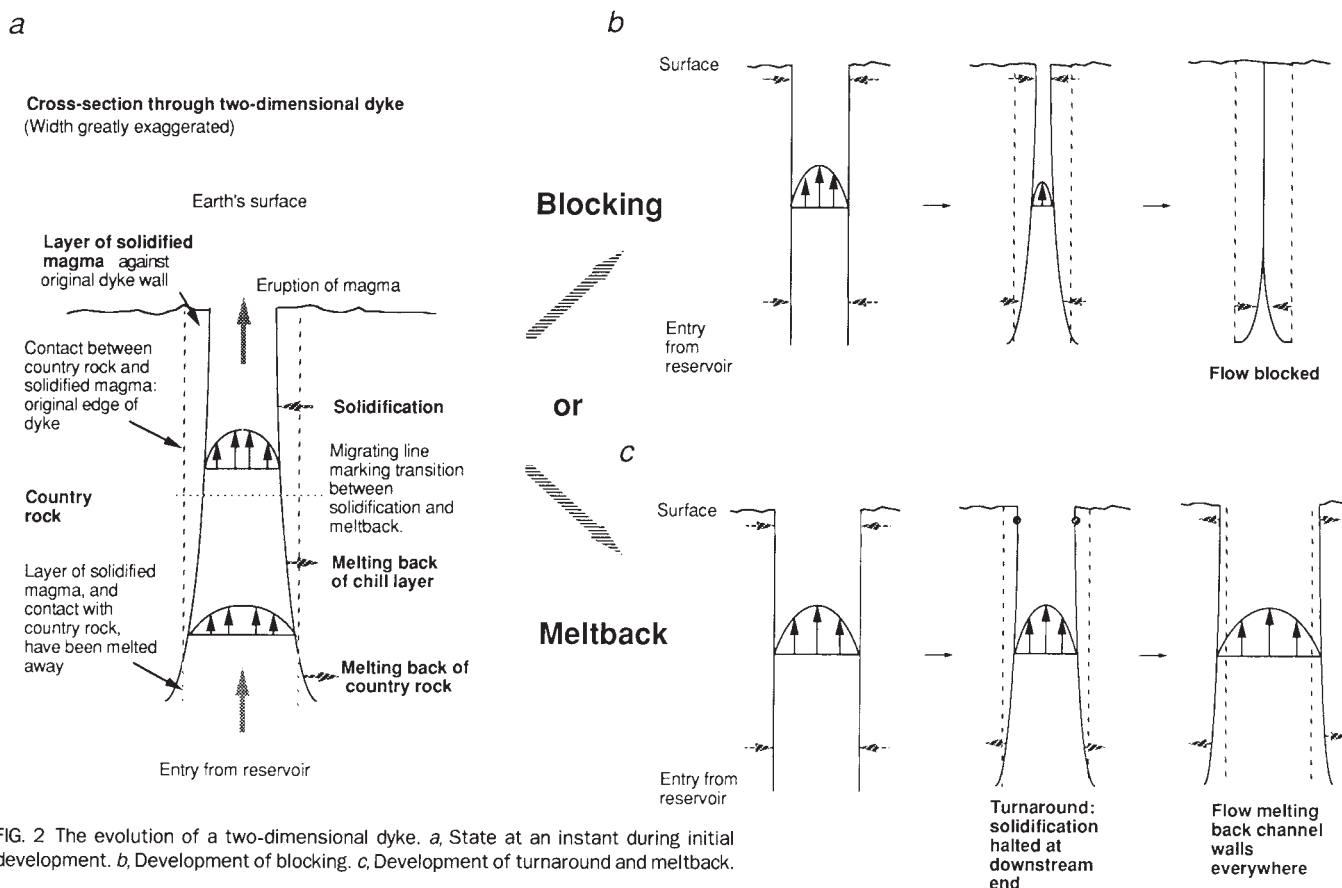


FIG. 2 The evolution of a two-dimensional dyke. *a*, State at an instant during initial development. *b*, Development of blocking. *c*, Development of turnaround and meltback.

result of melting is broadly consistent with some previous observations. Both the evolution of the Pu'u O'o series of en echelon dyke fissures in Hawaii (L. Wilson, personal communication) and the final width of 25 m for some fissure dykes in Iceland (M. R. Ryan, personal communication) seem to be

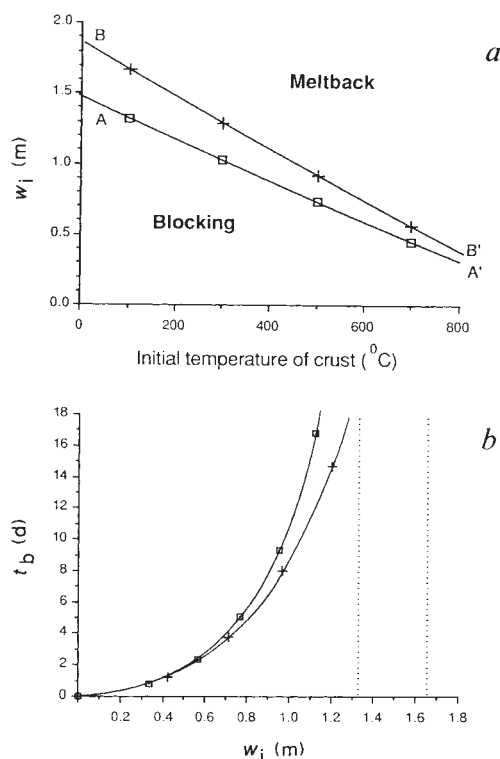


FIG. 3 Numerical results of model, for parameter values given in text. *a*, Regions of parameter space in which blocking or meltback are predicted for dyke length 2 km (squares) and 5 km (crosses). *b*, Time for dyke to become blocked as a function of initial dyke width, for $T_{\infty}=100^{\circ}\text{C}$. The time to block becomes infinite for widths approaching the critical value for meltback to occur (dotted lines (left) $H=2\text{ km}$; dotted lines (right) $H=5\text{ km}$). Symbols as in *a*.

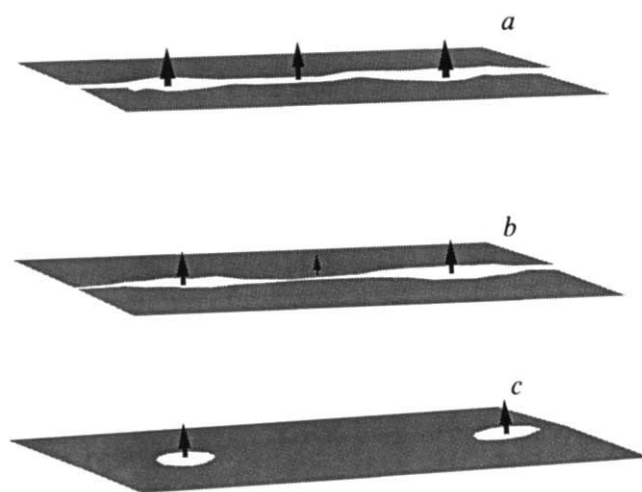


FIG. 4 A schematic (not to scale) diagram of flow localization and development of a surface fissure over time. *a*, initial irregularities of width (which may be small). *b*, faster solidification in narrow regions exaggerates the difference. *c*, flow through isolated vents melts back walls and flow is blocked elsewhere.

explicable only if melting back of the original dyke walls is postulated.

Further, the two-dimensional analysis quantifies an important feedback mechanism. Decreasing the width of the dyke restricts the flow. This reduces the advected heat supply which leads to further solidification and reduction of the width. Correspondingly, widening the dyke by melting the walls increases the advected heat supply which results in further melting. This feedback mechanism has an important role in a more realistic three-dimensional model, in which the width of the dyke varies, as depicted in Fig. 4. Cross-flows between sections of different initial widths will form because flow through the narrow sections will encounter greater constrictions because of preferential solidification. The cross-flows further reduce the advected heat flux in the downstream portions of the narrow sections and increase it at the wider portions, thereby exaggerating initially small differences in the width of the dyke (Fig. 4*b*). Our model predicts that, as a result of this mechanism, many fissure eruptions of moderate size are localized at isolated vents over a period of days (Fig. 4*c*), which is consistent with observations on Hawaii and elsewhere (refs 2, 3, 14 and T. L. Wright, personal communication) and with a previous qualitative discussion¹⁵. Eruptions from these vents will persist until limited by other volcanic processes. Sustained eruptions result in the formation of cylindrical conduits that are much wider than the original dyke and are preserved as volcanic plugs. □

Received 6 June; accepted 7 November 1989.

- Macdonald, G. A. & Abbot, A. T. *Volcanoes in the Sea* (University of Hawaii Press, Honolulu 1970).
- Richter, D. H., Eaton, J. P., Murata, K. J., Ault, W. A. & Krivoy, H. L. *U.S. Geol. Surv. Prof. Pap.* 537-E (1970).
- Thorarinnsson, S., Steinthorsson, S., Einarsson, T. H., Kristmannsdottir, H. & Oskarsson, N. *Nature* **241**, 372-375 (1973).
- Thorarinnsson, S. *Bull. Vol.* **33**, 910-929 (1969).
- Jaeger, J. C. *Basalts: The Poldervaart Treatise on Rocks of Basaltic Composition* (eds Hess, H. H. & Poldervaart, A.) 503-536 (Interscience, New York, 1968).
- Spence, D. A. & Turcotte, D. L. *J. geophys. Res.* **90**, 575-580 (1985).
- Lister, J. R. *J. Fluid Mech.* **210**, 263-280 (1990).
- Delaney, P. T. & Pollard, D. D. *Am. J. Sci.* **282**, 856-885 (1982).
- Delaney, P. T. & Pollard, D. D. *U.S. Geol. S. Prof. Pap.* 1202 (1981).
- Bruce, P. M. thesis, Univ. of Cambridge (1989).
- Bruce, P. M. & Huppert, H. E. in *Magma Transport and Storage* (ed. Ryan, M. P.) (Wiley, London, in the press).
- Huppert, H. E. & Sparks, R. S. J. *Contr. Miner. Petrol.* **75**, 279-289 (1980).
- Wilson, L. & Head, J. W. *J. geophys. Res.* **86**, 2971-3001 (1981).
- Anderson, I. *New Scientist* **1567**, 50-54 (1987).
- McBirney, A. R. *Igneous Petrology* (Freeman-Cooper, San Francisco, 1984).

ACKNOWLEDGEMENTS. We thank R. C. Kerr, J. R. Lister, R. S. J. Sparks, A. W. Woods and M. G. Worster for helpful discussions and P. T. Delaney, A. R. McBirney, D. D. Pollard, L. Wilson and T. L. Wright for stimulating comments. P.M.B. was supported by NERC studentship and H.E.H. was supported by the B.P. Venture Research Unit.

Hydrothermal plumes along the North Fiji Basin spreading axis

Yukihiro Nojiri*, Jun-ichiro Ishibashi†‡, Takayoshi Kawai*, Akira Otsuki*§ & Hitoshi Sakai†

* National Institute for Environmental Studies, PO Tsukuba, Ibaraki 305, Japan

† Ocean Research Institute, University of Tokyo, Nakano-ku, Tokyo 164, Japan

STEADY-STATE emanations of hydrothermal vent fluid along the mid-ocean ridges give rise to plumes rising 100-300 m above the sea floor, identified by anomalies of temperature, conductivity, chemistry and particles. Recently, a quite different type of hydrothermal plume created by a brief but massive release of high-temperature hydrothermal fluids was discovered by Baker *et al.* at a Juan de Fuca vent field¹⁻³. This gigantic thermal plume was termed a 'megaplume'. In December 1987, we found a large

‡ Present address: Faculty of Science, University of Tokyo, Bunkyo-ku, Tokyo 113, Japan.

§ Present address: Tokyo University of Fisheries, Minato-ku, Tokyo 108, Japan.

hydrothermal plume showing anomalies of temperature, methane and manganese above the North Fiji Basin spreading axis. The transience of this plume at $173^{\circ} 30' \text{ E}$, $18^{\circ} 50' \text{ S}$ was confirmed by our return visit to the site in November 1988. The discovery of megaplumes at two sites in very different tectonic environments indicates that this type of hydrothermal activity may make an important contribution to thermal and chemical fluxes from the mid-ocean ridges.

The North Fiji Basin is a 2,000–4,000-m-deep back-arc basin behind the active subduction zone of the Vanuatu Arc (Fig. 1a) with a spreading rate⁴ of $\sim 7.2 \text{ cm yr}^{-1}$. In December 1987, the 'Kaiyo 87' cruise was conducted (the STARMER project⁵) to locate the recent spreading axis more precisely⁶. The cruise followed up the results of a previous survey⁴ (SEAPSO leg 3, December 1985). Here we describe the results of conductivity–

temperature–depth (CTD) profiling and water sampling and discuss the nature of the plumes observed.

Because no precise vent locations had been discovered, we used eleven hydrocast stations along the ridge crest (Fig. 1b). The background concentrations of Mn and CH_4 ranged from 0.4 to 1.0 nmol kg^{-1} and from 0.2 to 0.3 nmol kg^{-1} , respectively. These are comparable with the background levels for deep waters of the open oceans^{7,8}. We observed Mn anomalies at all stations; the Mn concentration increased with depth below 1,900 or 2,200 m. The Mn concentrations of the bottom layer at the nine stations, located between $16^{\circ} 54' \text{ S}$ and $18^{\circ} 50' \text{ S}$, ranged between 5 and 10 nmol kg^{-1} . This indicates that the hydrothermal supply of Mn extends over a large area similar to the Juan de Fuca Ridge⁹. The anomalies of Mn at the northernmost and the southernmost stations (stations 11 and 15) were the smallest,

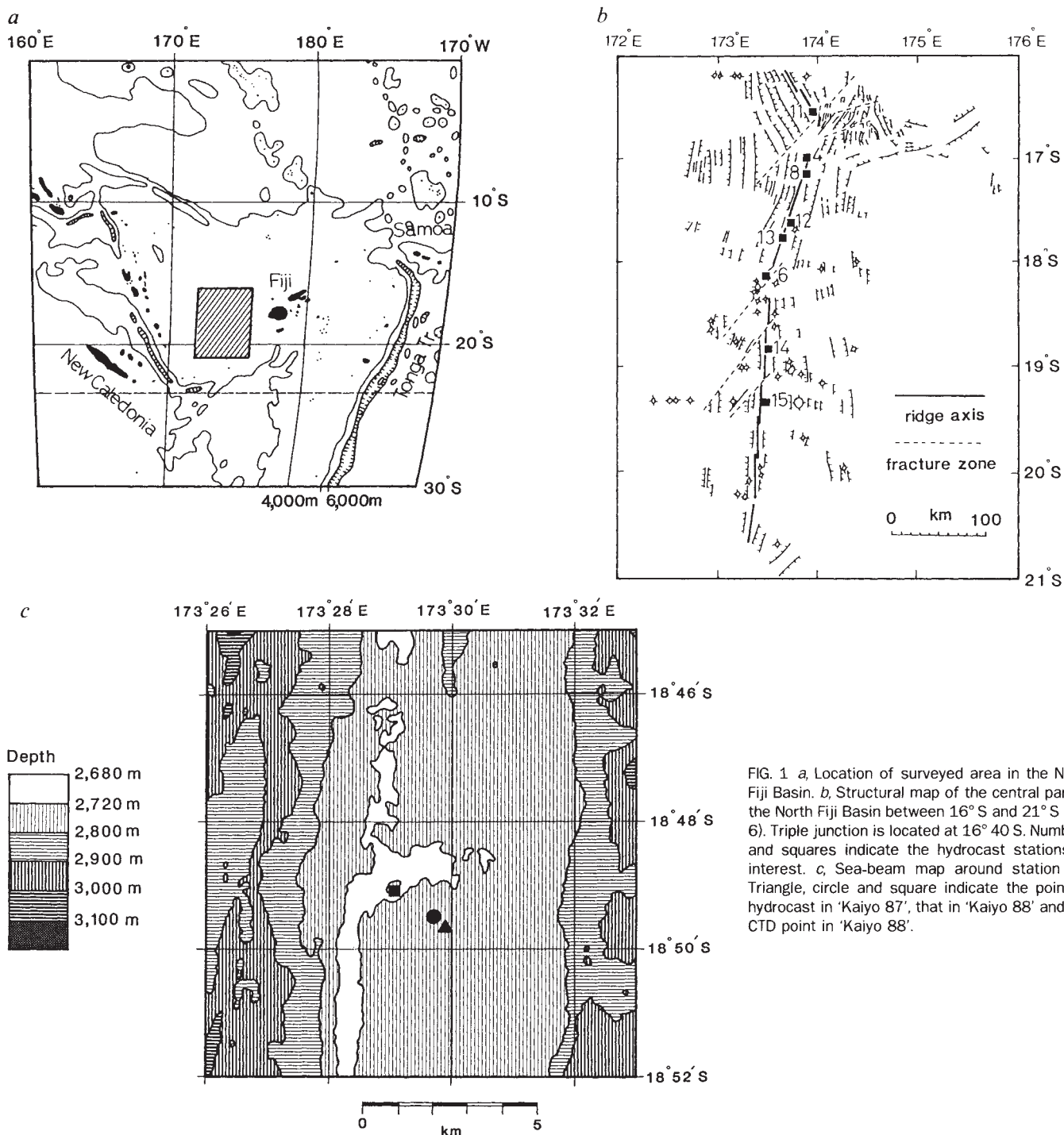


FIG. 1 a, Location of surveyed area in the North Fiji Basin. b, Structural map of the central part of the North Fiji Basin between 16° S and 21° S (ref. 6). Triple junction is located at $16^{\circ} 40' \text{ S}$. Numbers and squares indicate the hydrocast stations of interest. c, Sea-beam map around station 14. Triangle, circle and square indicate the point of hydrocast in 'Kaiyo 87', that in 'Kaiyo 88' and the CTD point in 'Kaiyo 88'.

TABLE 1 Estimated flux ratios of heat, Mn and CH₄ in hydrothermal solutions

Locations	Mn/heat (nmol cal ⁻¹)	CH ₄ /heat (nmol cal ⁻¹)	CH ₄ /Mn (mol mol ⁻¹)	References
EPR, 21° N, vent water	2–2.8	—	—	11
EPR, 20° S, VULCAN, large plume	0.3	—	—	7
EPR, 20° S, VULCAN, small plume	>1	—	—	7
Southern JF, plume	—	—	0.09	12
JF, SSS, vent water	7.6–15.7	—	—	13
JF, SSS, plume	1.4	—	—	14
JF, SSS, megaplume	0.37	—	—	2
JF, SSS, steady-state plume	3.5	—	—	2
JF, Endeavour, plume	0.4–1.0	6	6–16	15
Okinawa Trough, Iheya Glaben	—	—	0.6–3.2	16
NFB, SEAPSO, average deep water	—	—	0.12	4
NFB, SEAPSO, high CH ₄ plume	—	—	0.24–0.31	4
NFB, Kaiyo 87, high Mn plume	0.92	0.05	0.056	This work
NFB, Kaiyo 87, average deep water	—	—	0.121	This work
NFB, Kaiyo 87, high CH ₄ plume	0.13	0.05	0.386	This work

EPR, East Pacific Rise; JF, Juan de Fuca; NFB, North Fiji Basin.
SSS, southern symmetrical segment.

the maximum concentrations being <5 nmol kg⁻¹.

The plot of potential temperature θ against salinity S for the station 14 (triangle in Fig. 1c, 173° 30' E, 18° 50' S, 29 December 1987) data is shown in Fig. 2a. The apparent distortion reflects the θ -anomaly that is due to the injection of hydrothermal water into the ambient sea water. Figure 2b is the calculated profile of the excess temperature. It shows a large plume, having a sharp peak 800 m above the sea floor and a broad peak 650 m above the sea floor. There is another deeper plume 90 m above the sea floor. The density–depth profile shows no inversion so these are neutrally buoyant plumes spreading along the isodensity surface. The physical characteristics of the Juan de Fuca Ridge megaplume^{2,3} are as follows: to more than 800 m above the sea floor it is 700 m thick, has a 20-km-diameter spheroidal shape and a maximum θ -anomaly of 0.28 °C. Beneath the megaplume, a separate deeper plume was observed. The vertical scale and the θ -anomaly of the upper plume of station 14 are similar to those of the Juan de Fuca Ridge megaplume. The deeper plume corresponds to the steady-state plume of the Juan de Fuca Ridge. Smaller θ -anomalies were observed at station 8 and 13. At the other stations, smooth θ - S relationships were obtained.

Figure 3 displays the Mn and CH₄ profiles at station 14. The upper large plume in 'Kaiyo 87' is characterized by a low Mn concentration but the deeper narrow plume shows a high Mn concentration. The shapes of the vertical profiles of Mn in both the Juan de Fuca Ridge and the North Fiji Basin are very similar but the Juan de Fuca Ridge plumes had concentrations about seven times higher than that of the North Fiji Basin.

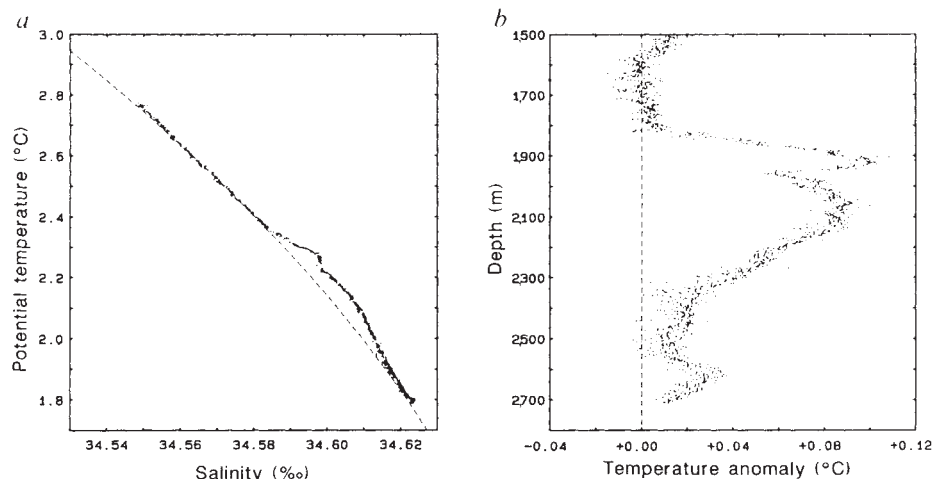
The correlation of Mn and CH₄ concentrations for station 14

is plotted in Fig. 4 along with those at station 12 in 'Kaiyo 87'. At station 12, we found no θ -anomaly but the concentration of Mn increased gradually with depth below 2,200 m. The linear CH₄/Mn relationship may represent an average injection ratio of CH₄ and Mn in the area. The slope (0.121 mol mol⁻¹, regression line B in Fig. 4) is the same as the averaged CH₄/Mn ratio (0.12 mol mol⁻¹) in the previous SEAPSO results⁴ (Table 1).

The available data for the hydrothermal-supply ratios of CH₄/heat, Mn/heat and CH₄/Mn are summarized in Table 1. Regression line A in Fig. 4 corresponds to the upper plume of station 14, which has a high CH₄/Mn ratio. The regression line C corresponds to the deeper plume of this station, which has a low CH₄/Mn ratio. We calculated the Mn/heat and CH₄/heat ratios for both plumes. The relationship of Mn/heat ratio between the upper and lower plumes of station 14 corresponds to that between the megaplume and the accompanying steady-state plume of the Juan de Fuca Ridge^{2,3}.

We made a return visit to station 14 on 20 November 1988 during the 'Kaiyo 88' cruise (circle in Fig. 1c). The distance between the two sampling points was only a few hundred metres, but we found no θ -anomaly. The Mn profile in 'Kaiyo 88' is shown in Fig. 3 (closed circles). The upper plume of the previous year had disappeared but a large Mn anomaly was observed near the sea floor. We detected a small θ -anomaly of 0.02 °C, with a profile similar to the lower plume of the previous year, at a nearby CTD station (square in Fig. 1c). We also detected vent animal communities (deep-tow camera observation) during 'Kaiyo 88'. These observations indicate hydrothermal activity in the area. It is likely that the 'Kaiyo 88' hydrocast detected

FIG. 2 a, Potential temperature (θ) plotted against salinity (S) at station 14 of 'Kaiyo 87'. Seabird model 9 CTD with rosette sampler was used. The resolution of the thermometer and conductivity cell were 0.0003 °C and 0.00004 S m⁻¹, respectively. Data obtained by CTD was averaged in every 1-m interval for water-column depths >1,500 m. The broken line is the smoothed curve obtained from the θ - S relationship at station 15, the nearest station having no temperature anomaly to station 14. b, Vertical profile of excess temperature in the water column of station 14 of 'Kaiyo 87', calculated as a difference between the potential temperatures at stations 14 and 15 at the same salinity.



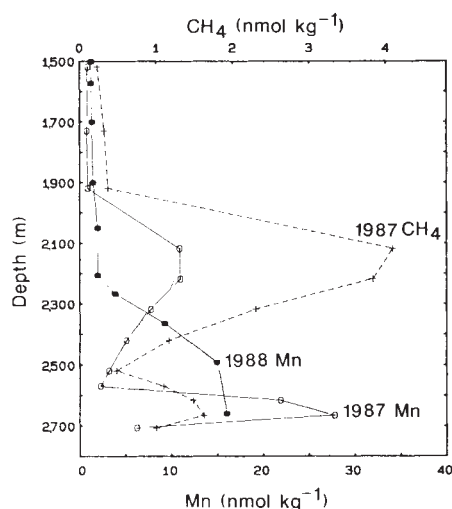


FIG. 3 a, Vertical profiles of Mn and CH_4 at station 14. Open circles, crosses and closed circles are Mn of 'Kaiyo 87', CH_4 of 'Kaiyo 87' and Mn of 'Kaiyo 88', respectively. Water samples were collected with Teflon-lined 8 l Go-Flo Niskin bottles mounted on rosette multi-sampler. Mn in sea water was collected aboard ship on a Nucleopore membrane filter with added ferric hydroxide precipitate. Mn in both dissolved and particulate forms are effectively separated from the seawater matrix. The precipitate was dissolved in hydrochloric acid and analysed by inductively-coupled plasma emission spectrometry on shore. Samples for CH_4 analysis were sealed in glass bottles with septum stopper after poisoning with HgCl_2 and preserved in a refrigerator until they were analysed on shore¹⁶. The limit of determination of Mn and CH_4 analyses were both 0.1 nmol kg^{-1} . The reproducibility of Mn analysis was 2% at 7 nmol kg^{-1} , but 10% at 1 nmol kg^{-1} . CH_4 analysis was accurate within 4% at 1 nmol kg^{-1} .

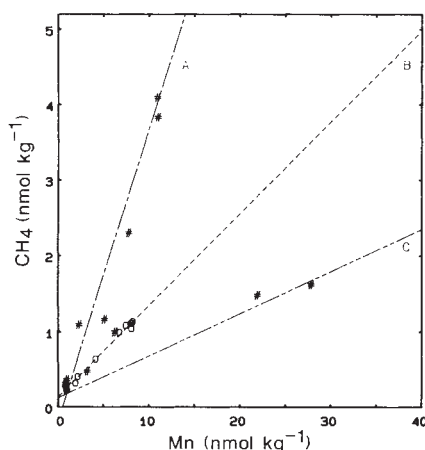


FIG. 4 Correlation between Mn and CH_4 concentrations in sampled sea water of stations 12(o) and 14(#), of 'Kaiyo 87' cruise. Regression lines were calculated for the data of stations 8, 12, 13 and 14. A, $\text{CH}_4(\text{nmol kg}^{-1}) = 0.386 \times \text{Mn}(\text{nmol kg}^{-1}) - 0.16$ (correlation coefficient, $r = 0.90$, number of points, $n = 9$); B, $\text{CH}_4(\text{nmol kg}^{-1}) = 0.121 \times \text{Mn}(\text{nmol kg}^{-1}) + 0.14$ ($r = 0.99$, $n = 11$); C, $\text{CH}_4(\text{nmol kg}^{-1}) = 0.056 \times \text{Mn}(\text{nmol kg}^{-1}) + 0.13$ ($r = 0.93$, $n = 9$).

the steady-state hydrothermal effluent emerging at the spreading axis and that the upper plume of the previous year had disappeared.

The identification of the upper plume of the 'Kaiyo 87' station 14 as a megaplume depends on its vertical scale, its large θ -anomaly, the low concentration of Mn and its absence in 'Kaiyo 88'. These characteristics are similar to those of the Juan de Fuca Ridge megaplume. During an earlier cruise, Craig *et al.*¹⁰ discovered a large plume having a very elevated concentration (13 nmol kg^{-1}) of CH_4 in the North Fiji Basin. Our hydrocast at the location (station 6 in Fig. 1b) revealed no plume-shaped anomaly of Mn and the CTD profile showed no

temperature anomaly. The CH_4 plume observed by Craig *et al.* may have been another megaplume.

The different Mn/heat ratio of the upper plume and lower plume of station 14 could be a consequence of a different origin. The extraction efficiency of Mn into vent fluid would depend on the reaction time of the fluid with hot basalt and the effective water/basalt ratio during plume formation. The sudden catastrophic production of a megaplume might thus be expected to result in a low Mn/heat ratio compared with a steady-state plume.

The observations of megaplume-type hydrothermal effluent at different ocean ridges indicates that this type of hydrothermal activity may be a common phenomenon in rift systems. The possibility of a frequent occurrence of megaplumes is significant for the global hydrothermal fluxes of heat and chemical components. More extensive investigations of megaplumes, especially of their frequency, spatial scale and chemical nature, are necessary. □

Received 1 April; accepted 19 October 1989.

1. Cann, J. & Strens, R. *Nature* **329**, 104 (1987).
2. Baker, E. T., Massoth, F. J. & Feely, R. A. *Nature* **329**, 149–151 (1987).
3. Lupton, J. E., Baker, E. T. & Massoth, G. J. *Nature* **337**, 161–164 (1989).
4. Auzende, J. M. *et al. Tectonophysics* **146**, 317–351 (1988).
5. Swinbanks, D. *Nature* **328**, 4 (1987).
6. Auzende, J. M. *et al. C. r. heb. Séanc. Acad. Sci., Paris, Ser. 2* **306**, 971–978 (1988).
7. Klinkhammer, G. & Hudson, A. *Earth planet. Sci. Lett.* **79**, 241–249 (1986).
8. Charlou, J. L., Rona, P. & Bougault, H. *J. mar. Res.* **45**, 461–472 (1987).
9. Jones, C. J., Johnson, H. P. & Delaney, J. R. *Geophys. Res. Lett.* **8**, 873–876 (1981).
10. Craig, H., Craig, V. K. & Kim, K. R. *Eos* **68**, 100–101 (1987).
11. Von Damm, K. L. *et al. Geochim. cosmochim. Acta* **49**, 2197–2220 (1985).
12. Winn, C. D., Karl, D. M. & Massoth, G. J. *Nature* **320**, 744–746 (1986).
13. Von Damm, K. L. & Bischoff, J. L. *J. geophys. Res.* **92**, 11334–11346 (1987).
14. Baker, T. E. & Massoth, G. J. *Science* **234**, 980–982 (1986).
15. Rosenberg, N. D. *et al. Nature* **334**, 604–607 (1988).
16. Ishibashi, J. *et al. Geochem. J.* **22**, 107–114 (1988).

ACKNOWLEDGMENTS. We thank E. Honza, J. M. Auzende and the other shipboard scientists, and the crew of the RV *Kaiyo*. We also thank T. Matsumoto and A. Tanaka. This study was funded by the Science and Technology Agency of Japan.

An unusual hopane biodegradation sequence in tar sands from the Pt Arena (Monterey) Formation

A. G. Requejo*† & H. I. Halpern*

* ARCO Oil and Gas Company, 2300 West Plano Parkway, Plano, Texas 75075, USA

HOPANES are ubiquitous in petroleum source rocks and crude oils. These compounds are thought to be derived mainly from a specific oxygenated precursor, bacteriohopanetetrol, found in the cell walls of prokaryotic organisms¹, and commonly occur as a homologous series of structurally related isomers containing between 27 and 35 carbon atoms; hopanes containing more than 35 carbon atoms have also been reported². Hopanes and other pentacyclic triterpanes are relatively resistant to microbial degradation^{3–5}, making them useful in correlating moderately biodegraded oils to their source rocks⁵. During the advanced stages of biodegradation, however, these compounds can be altered, being converted either to 25-normethyl homologues⁴ or to other, as yet uncharacterized, products. Selective removal of individual compounds may occur; for example, 22R extended homohopanes may be degraded whereas 22S epimers are preserved^{6–8}. Here we report an unusual biodegradation sequence of pentacyclic triterpanes in tar sands from the Pt Arena Formation, California, in which the 22S and 22R isomers of the 17 α (H), 21 β (H)-C₃₅ extended hopanes are selectively preserved. Demethylation at the C₂₅ position is observed for most hopanes in severely biodegraded samples, with

† Present address: Exxon Production Research Company, PO Box 2189, Houston, Texas 77001, USA.

the exception of the C_{35} extended hopanes. Two commonly occurring trisnorhopanes⁹, designated Ts ($18\alpha(H)$, $21\beta(H)$ -22, 29, 30-trisnorhopane) and Tm ($17\alpha(H)$, $21\beta(H)$ -22, 29, 30-trisnorhopane), also seem resistant to degradation. These findings, together with previous studies of crude-oil biodegradation, imply that biodegradative alteration of hopanes can proceed through numerous pathways, which can result in widely differing distributions.

We have analysed several tar-sand samples from the Miocene Pt Arena (Monterey) Formation in the Pt Arena Basin, California. Tar-stained sandstones are distributed throughout siliceous organic mudstones exposed at several locations along the northern California coastline. The thickest continuous tar-sand unit measures ~15 m, and consists of graded, medium-coarse-grained sands containing mudstone clasts at the base of some individual beds. We collected the samples near this unit, either directly from the exposure or from eroded fragments lying nearby. All samples seemed heavily weathered and varied in composition from a pliable tar distributed throughout a sand matrix to more bituminous, less viscous samples which would flow upon heating by sunlight. We assume that most, if not all, biodegradation in these tar sands occurred following their exposure.

Methylene chloride extracts of the tar sands consist of a sulphur-rich, bituminous material enriched in aromatic and polar components. The composition of C_{15+} solvent-extractable material in four samples is shown in Table 1. Saturated hydrocarbons make up a small fraction of the extract (3.7–11.1%) and the asphaltene contents vary from 16.5 to 54.1%. The sulphur content of the extractable material varies from 3.2 to 5.8% (Table 1). These characteristics are typical of high-sulphur petroleum, such as those in the Monterey Formation¹⁰, that have been severely biodegraded. Isotopically heavy ^{13}C compositions for the saturated and aromatic hydrocarbons (Table 1) indicate that the tar was most likely generated from facies equivalent to the Monterey Formation in the Pt Arena Basin^{10,11}.

The saturated hydrocarbons isolated from the four samples consist primarily of naphthenic compounds and are devoid of *n*-alkanes and isoprenoids, a composition characteristic of moderate to severe biodegradation^{5,12}. We confirmed the presence of tricyclic and pentacyclic triterpanes by gas chromatographic/mass spectrometric (GC/MS) analysis of the saturated-hydrocarbon fractions. These results also revealed that 'normal' steranes (those with methyl groups substituted at the C_{10} - and C_{13} positions) have been completely biodegraded in all samples, and only diasteranes (C_5 and C_{14} methyl substituted) can be identified at low abundance in $m/z = 217$ mass chromatograms. The 'normal' steranes are the most susceptible to degradation among tetra- and pentacyclic hydrocarbon biomarkers and their removal indicates that the extent of biodegradative weathering has progressed to at least level 7 on Volkman *et al.*'s scale⁴, which they describe as "very extensive" alteration.

By contrast to the steranes, hopanes in these samples have been altered to varying degrees. Figure 1 shows $m/z = 191$ mass chromatograms corresponding to the four samples in Table 1. The distribution in sample 87-2408 (Fig. 1a), typical of Monterey-derived oils, exhibits several characteristic features such as the occurrence of 28, 30-bisnorhopanes (labelled 28H in Fig. 1) and the C_{35} extended hopanes (labelled 35H, Fig. 1) in higher abundance than the C_{34} homologues^{13,14}. These features have been variously related to source rocks deposited in either a sulphur-rich environment, in the case of the bisnorhopanes^{14,15}, or hypersaline environments, in the case of the C_{35} extended hopanes^{16–18}. We characterize the distribution in this sample as 'unaltered', although we are not certain how closely it resembles the pristine distribution before any biodegradation. Comparison to triterpane distributions of undegraded Monterey-derived oils^{13,14} suggests that the hopanes in this sample have not undergone significant alteration.

Figure 1b, c show the effects of increasing biodegradation on

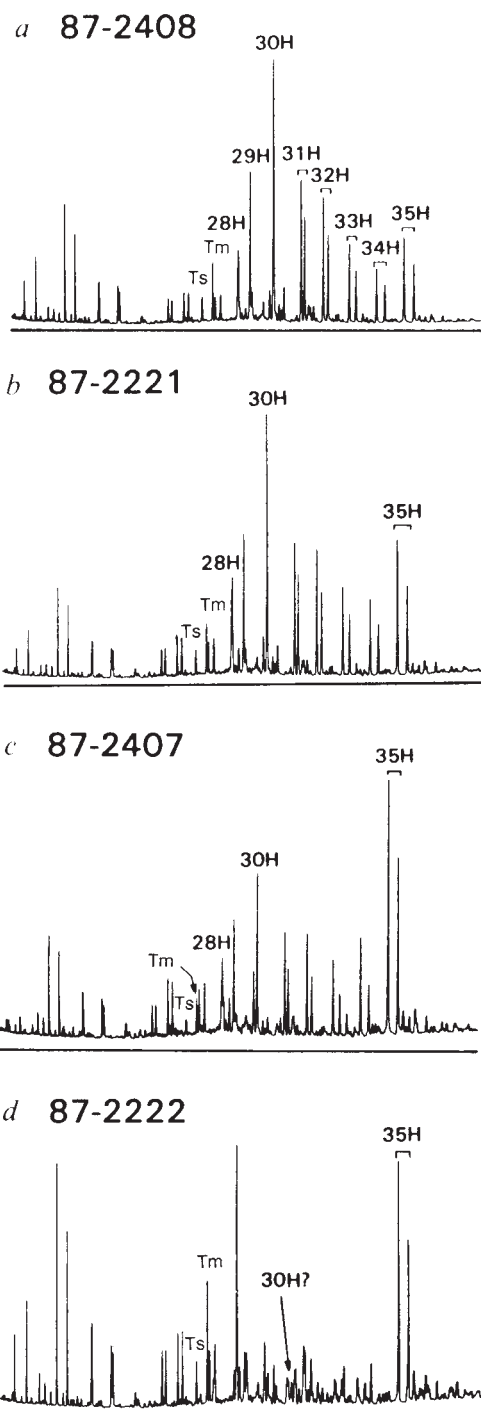


FIG. 1 Mass chromatograms of $m/z = 191$ depicting triterpane distributions in Pt Arena tar-sand samples. Homologous series of hopanes possessing $17\alpha(H)$, $21\beta(H)$ -stereochemistry are denoted by their carbon number followed by the letter H. The two commonly occurring trisnorhopanes are designated Ts and Tm. Note the gradual loss of most hopanes relative to the C_{35} compounds with increasing degree of alteration (proceeding from a to d). In the most severely altered sample (Fig. 1d) the principal triterpenoid compounds consist of tricyclic terpanes (not labelled), the trisnorhopanes Ts and Tm, the C_{35} hopanes, and various unidentified components. Bisnorhopane was not detected in this sample. Calculated concentrations of several hopanes (Table 1) indicate that the C_{35} compounds are being selectively preserved. Chromatographic conditions: 25-m fused-silica capillary column containing 5% phenyl methyl silicone (0.20-mm inner diameter, 0.33- μ m film thickness) temperature programmed from 60 to 220°C at 25°C min⁻¹ and 220 to 325°C at 2°C min⁻¹ after a 1.0-min isothermal hold, helium carrier gas (1.0 cm³ min⁻¹); injector temperature 275°C; mass spectrometric analysis carried out in SIM (selected ion monitoring) mode using a Hewlett-Packard 5970 Mass Selective Detector.

TABLE 1 Composition of extractable organic matter and calculated values for various hopane parameters in Pt Arena tar sands

Sample no.	Extractable organic matter							Triterpane parameters							
	Content (p.p.m. dry weight)	Sulphur (weight %)	SATS	Composition*			¹³ C†		Hopane ratios‡			Concentrations§			
				AROS	HETS	ASPH	SATS	AROS	28/30	35/30	%22S	Ts/Tm	35H	30H	28H
87-2408	30,000	3.2	11.1	45.2	26.2	16.5	-23.3	-22.7	0.55	0.90	57	0.45	438	485	265
87-2221	72,000	4.2	10.3	42.2	26.2	19.9	-23.3	-22.7	0.82	1.59	58	0.50	452	284	233
87-2407	126,000	5.8	3.7	20.1	21.6	54.1	-23.5	-22.7	0.94	4.99	59	0.50	369	74	70
87-2222	76,000	4.5	9.2	47.2	20.2	22.2	-23.3	-22.8	NC	14.49	62	0.37	461	32	NC

NC=Not calculated.

* SATS, AROS, HETS, ASPH=saturated and aromatic hydrocarbons, heterocompounds and asphaltenes, respectively.

 † ^{13}C composition reported as $\delta^{13}\text{C}=[(^{13}\text{C}/^{12}\text{C})_{\text{sample}}/(^{13}\text{C}/^{12}\text{C})_{\text{standard}}]-1$. ‡ 28/30=28,30-bisnorhopanes/17 α (H), 21 β (H)-hopane; 35/30= C_{35} extended hopanes (22S+22R)/17 α (H), 21 β (H)-hopane; %22S= C_{31} extended hopanes (22S)/(22S+22R); T/T=17 α (H)-trisorneohopane/18 α (H)-trisorneohopane. § 28H, 30H, 35H=Concentrations of 28, 30-bisnorhopanes, 17 α (H), 21 β (H)-hopane and C_{35} extended hopanes (22S+22R), respectively, expressed per unit pentane-soluble extractable material. Concentrations were calculated relative to an added internal standard (5 α (H)-androstane) and are corrected for differences in gas chromatograph/mass spectrometer response.

the hopanes. These include a preferential decrease in the abundance of 17 α (H), 21 β (H)-hopane (labelled 30H, Fig. 1) relative to other hopanes and a slight decrease in the abundance of the 22R isomer of the C_{33} and C_{34} extended hopanes relative to the 22S isomer. Most striking, however, is the increasing abundance of the 22S and 22R isomers of the C_{35} hopanes relative to other hopanes. This effect can be seen at its incipient stages in sample 87-2407 (Fig. 1c). In the most severely altered sample (sample 87-2222), hopanes containing between 28 and 34 carbons have either disappeared or are present in exceedingly low abundance, and the C_{35} compounds are prevalent (Fig. 1d). The two most commonly occurring trisnorhopanes, designated Ts and Tm⁹, also seem to remain relatively unaltered throughout the entire biodegradation sequence.

The trends shown in Fig. 1 are quantified in Table 1. The calculated bisnorhopane/hopane ratio shows an increase (from 0.55 to 0.94), suggesting that hopane is degraded slightly more rapidly during the incipient stages of biodegradation. Similarly, the 22S% values (calculated from C_{31} extended hopanes) increase very slightly with increasing extent of alteration, although we consider the value calculated for sample 87-2222 uncertain because of the difficulty in unambiguously identifying these compounds at the advanced stages of hopane alteration evident in this sample. By comparison, the ratio of $\text{C}_{35}/\text{C}_{30}$ hopanes increases progressively from 0.90 to 14.49 with increasing extent of biodegradative alteration (Table 1). The Ts/Tm ratio changes little in the four samples.

In the most degraded samples we can document the formation of 25-normethyl homologues. Figure 2 shows mass chromatograms of $m/z=177$ and 191 for a severely altered sample, 87-947. The $m/z=177$ ion is diagnostic for triterpanes demethylated in the A/B ring system⁴. Overall, the distribution depicted in the $m/z=177$ chromatogram is similar to the 'unaltered' hopane distribution (Fig. 1a), but differs in several respects. We attribute some of these differences to the fact that the response at $m/z=177$ for 25-normethyl hopanes is not the same as that for the corresponding hopane at $m/z=191$. Thus the greater abundance of the demethylated C_{28} compound in the $m/z=177$ chromatogram relative to the C_{29} hopane in the $m/z=191$ chromatogram is due to differences in the mass spectra of the respective compounds. Two characteristics of the respective distributions, however, are truly different. First, the demethylated C_{34} homologues of the C_{35} hopanes occur at much lower abundance than in the 'unaltered' hopane distribution. Second, the demethylated C_{26} homologues of the trisnorhopanes Ts and Tm are absent. The latter result differs from studies of biodegraded Australian oils, in which demethylation of these compounds with a distribution similar to that of the hopanes was observed⁴. Both observations indicate selective preservation of these compounds in severely altered samples after most other

hopanes have been degraded (such as in Fig. 1d). These features are illustrated in Fig. 2.

Several explanations can be invoked to account for the preservation of the C_{35} extended hopanes. The carbon skeleton of these compounds very closely resembles that of the bacteriohopanetetrol precursor. Because bacteria mediate the degradation process, it might be suggested that these compounds were added to the bituminous material in the tar sands during the biodegradation process through the incorporation of microbial remains. We can dismiss this argument for several reasons. Absolute biomarker concentrations, calculated relative to an added internal standard, indicate that the concentration of the C_{35} compounds is relatively invariant from the least to the most severely degraded sample. Whereas concentrations of bisnorhopane and hopane shown in Table 1 decrease with increasing degree of alteration, concentrations of C_{35} extended hopanes remain relatively constant at 369–461 p.p.m. The lowest C_{35} -hopane concentration is associated with the one sample having a significantly higher asphaltene content (Table 1), which might account for the slightly diminished value. This constancy would

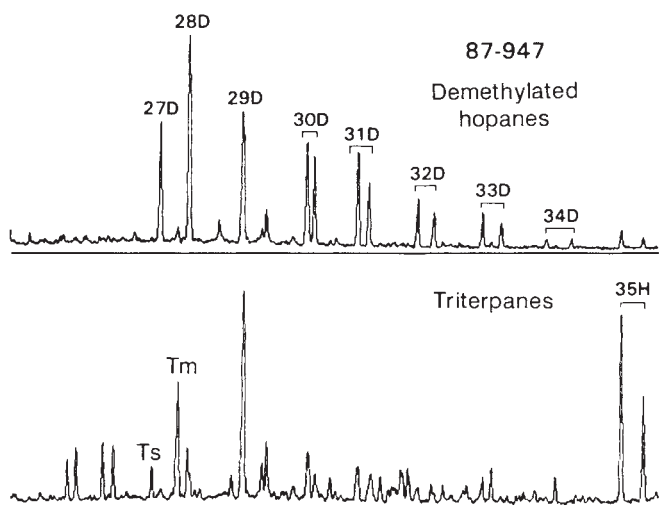


FIG. 2 Mass chromatograms of $m/z=177$ and $m/z=191$ depicting distributions of demethylated hopanes and hopanes, respectively, in a severely altered tar-sand sample (87-947). Notation as in Fig. 1; demethylated hopanes are denoted by the letter D. Demethylated homologues for most hopanes shown in Fig. 1a are present in the $m/z=177$ chromatogram with the exception of the trisnorhopanes Ts and Tm and the C_{35} hopanes. These compounds are the only hopanes recognized in the $m/z=191$ chromatogram. Chromatographic conditions as in Fig. 1; mass spectrometric analysis conducted using a Finnigan 8222 magnetic/electric-sector instrument (1,000 resolution).

be more characteristic of selective preservation, as it would be very unusual for these compounds to be added at the same rate as they are removed, and also rules out mechanisms involving progressive degradation of the C_{35} compounds at a greatly diminished rate relative to the other hopanes. Further, it would be highly unusual for these compounds to come from a biological source in the geologically preferred distribution, nearly completely isomerized at the C_{22} position (Fig. 1d).

The most plausible explanation for these observations is that, under the surface weathering conditions to which these tar sands have been exposed, the C_{35} extended hopanes are resistant to biodegradation. The same conclusion would seem to apply for the trisnorhopanes Ts and Tm. To the best of our knowledge, this hopane biodegradation sequence has not been reported previously. It is therefore difficult to assess the extent of its occurrence in the alteration of triterpanes in biodegraded oils. Enhanced C_{35} -hopane abundances are characteristic of specific environments of deposition, hence these observations may only apply to biodegradation of oils generated from source rocks such as carbonates, hypersaline sources or Monterey Formation analogues. Our observations are, however, somewhat serendipitous because they came about as a result of analysis of various tar-sand samples from the same location. Similar degradation sequences might go undetected unless several samples of the same oil that were biodegraded to varying extents were analysed.

These findings, together with previous studies of crude-oil biodegradation, imply that biodegradative alteration of hopanes can proceed through numerous pathways, which can result in widely differing distributions. Collectively, these pathways are probably more complex than the generalized degradation sequences described by either Volkman *et al.*⁴ or Connan⁵. The environmental conditions under which biodegradation takes

place (such as the temperature, oxygen level or salinity associated with in-reservoir conditions) may govern the composition of microbial communities active in biodegradative processes⁷, which in turn dictate the specific molecular transformations that are observed. This has been substantiated by results of various laboratory simulation studies^{8,19}. Caution must be used when evaluating source-rock deposition environments from analyses of hopane distributions in oils^{16,18}, particularly if the assessments involve analysis of petroleum seepage²⁰, as the alteration effects associated with biodegradation can be severe. □

Received 6 October; accepted 17 October 1989.

1. Ourisson, G., Albrecht, P. & Rohmer, M. *Pure appl. Sci.* **10**, 709–729 (1979).
2. Rulkötter, J. & Philp, P. *Nature* **292**, 616–618 (1982).
3. Seifert, W. K. & Moldowan, J. M. *Geochim. cosmochim. Acta* **43**, 111–126 (1979).
4. Volkman, J. K., Alexander, R., Kagi, R. I. & Woodhouse, G. W. *Geochim. cosmochim. Acta* **47**, 785–794 (1983).
5. Connan, J. in *Advances in Petroleum Geochemistry* Vol. 1 (eds Brooks, J. & Welte, D.) 299–335 (Academic, London, 1984).
6. Hoffman, C. F. & Strausz, O. P. *Am. Ass. Petrol. Geol.* **70**, 1113–1128, 1986.
7. Seifert, W. K., Moldowan, J. M. & Demaison, D. J. *Org. Geochem.* **6**, 633–643 (1984).
8. Goodwin, N. S., Park, P. J. D. & Rawlinson, T. in *Advances in Organic Geochemistry 1981* (eds Bjoroy, M. *et al.*) 650–658 (Wiley, Chichester, 1983).
9. Seifert, W. K. & Moldowan, J. M. *Geochim. cosmochim. Acta* **42**, 77–95 (1978).
10. Orr, W. I. *Org. Geochem.* **10**, 499–516 (1986).
11. Magoon, L. & Isaacs, C. M. in *Petroleum Generation and Occurrence in the Miocene Monterey Formation, California* (ed. Isaacs, C. M. *et al.*) 201–211 (Soc. Econ. Paleontologists Mineralogists, 1983).
12. Bailey, N. J. M., Jobson, A. M. & Rogers, M. A. *Chem. Geol.* **11**, 203–211 (1973).
13. Sofer, Z. *Org. Geochem.* **13**, 939–945 (1988).
14. Curiale, J. A., Cameron, D. & Davis, D. V. *Geochim. cosmochim. Acta* **49**, 271–288 (1985).
15. Grantham, P. J., Posthuma, J. & de Groot, K. in *Advances in Organic Geochemistry 1979* (eds Douglas, A. G. & Maxwell, J. R.) 29–38 (Pergamon, Oxford, 1980).
16. Connan, J. & Dessort, D. *Org. Geochem.* **11**, 103–113 (1987).
17. ten Haven, H. L., de Leeuw, J. W. & Schenck, P. A. *Geochim. cosmochim. Acta* **49**, 2181–2191 (1985).
18. ten Haven, H. L. *et al.* in *Lacustrine Petroleum Source Rocks, Geol. Soc. Spec. Publ.* **40** (eds Fleet, A. J. *et al.*) 123–140 (Blackwell Scientific, Oxford, 1988).
19. Rubinstein, I., Strausz, O. P., Spyckerelle, C., Crawford, R. J. & Westlake, D. W. S. *Geochim. cosmochim. Acta* **41**, 1341–1353 (1977).
20. Kennicutt, M. C. & Brooks, J. M. *Appl. Geochem.* **3**, 573–582 (1988).

Bacterial mats from Crater Lake, Oregon and their relationship to possible deep-lake hydrothermal venting

Jack Dymond, Robert W. Collier* & Mary E. Watwood†

* College of Oceanography, Oregon State University, Corvallis, Oregon 97331, USA

† Department of Biology, The University of New Mexico, Albuquerque, New Mexico 87131, USA

CRATER Lake is located in a caldera on Mt Mazama, a volcanic centre in the Oregon Cascades which has been active for more than 400,000 years¹. The 594-m-deep lake is a consequence of a climactic eruption which occurred $6,845 \pm 50$ years ago; however, caldera volcanism took place as recently as 4,000 years ago¹. It has been suggested that some of the physical and chemical features of the lake result from hydrothermal inputs^{2–7}. Here we present submersible observations of the bottom of Crater Lake, which reveal communities of bacteria that are usually associated with anomalously warm, saline waters. The bacteria seem to derive energy from the oxidation of ferrous iron to fuel their metabolism. We propose that the mats are indicators of diffuse hydrothermal venting into the deep lake.

During August of 1988, 16 dives in the one-person submersible *Deep Rover* were carried out in the lake. Nine dives were in the south-central part of the lake, an area of high heat flow⁴ and anomalous compositions and temperatures of the bottom water⁷. We observed bacterial mats at six locations in this area; at some sites multiple mats were present. All mats observed were within 300 m of each other and were either on or near an escarpment with a north-east/south-west orientation. The mats are

gelatinous, low-density, 15–300 cm in diameter and 2–15 cm thick. Some of the largest ones are draped over steep rock outcrops (Fig. 1), and the smaller mats lie directly on sediments. Sediments near the mats are dark and contrast sharply with the normal buff-coloured sediments. The maximum water temperature anomaly in the immediate vicinity of the mats was 0.1 °C. Insertion of a temperature probe into the mat, however, revealed temperatures ranging from 4.70 to 9.52 °C, up to 6.0 °C above the typical 3.5 °C bottom water. In addition to the

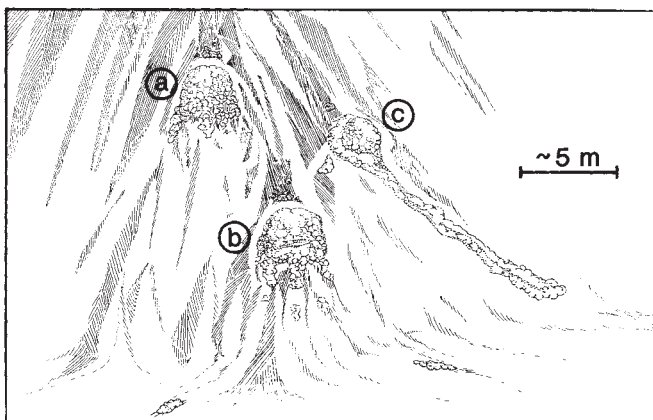


FIG. 1 A schematic drawing of bacterial mats observed during dive 179. The three principal mats observed at this site are identified by 'a', 'b', 'c'. These large mats seem to be attached to rocky cliff faces. Mat 'c' extends down the cliff face a further 10 m below the primary structure. This extension seems to be actively growing and not simply a consequence of pieces that roll down the slope. Small mats are growing at the base of the slope on sediments. The sediments along the base of the slope are discoloured, apparently because of Fe and Mn precipitation. Mat sample and mat waters shown in Tables 1 and 2 were recovered from 'a'.

TABLE 1 The composition of the bacterial mat and other iron-rich deposits

	Bacterial mat	Galapagos mounds ²⁶	Santorini ²⁷	Drive 187 crust*
Al	0.21	0.32	0.87	0.46
Si	7.2	13.3	6.8	11.7
P	3.7	—	0.11	1.96
Ca	0.8	1.4	1.3	0.87
Fe	37.9	27.5	27.8	35.5
Mn	0.014	7.9	0.045	0.45
As	0.36	—	—	0.08
Ba	0.02	0.07	0.005	0.05
S	0.3	—	0.41	—
C	2.44	—	4.0	—
N	0.97	—	—	—

All values are in weight per cent as determined by complete dissolution with HF/HNO₃ acids and analysis by atomic absorption spectrophotometry. The dashes indicate that no measurements were made.

* A lithified crust recovered from a sedimented bottom in the general region of the bacterial mats.

temperature anomalies associated with the mats, waters with anomalous concentrations of ³He, Mn, ²²²Rn, Li, Si and Cl have been observed at and near these features⁸⁻¹⁰.

A predominantly bacterial composition for the mats (genera *Gallionella* and *Leptothrix*) is confirmed by scanning electron microscopy (Fig. 2). The sheath-forming *Gallionella* is generally considered to be a chemolithoautotroph¹¹ as it requires a source of reduced iron¹² and will grow in the absence of significant organic carbon¹³. It is not necessarily thermophilic because it has been found in low-temperature springs, wells and drainages¹², but this genus has been observed in thermal springs with temperatures up to 47 °C (ref. 14). Because of the accumulations of ferric hydroxide and manganese dioxide accumulations on the sheaths of *Leptothrix*, these bacteria are thought to oxidize ferrous and manganous ions¹⁵. An autotrophic metabolism for *Leptothrix* has not been demonstrated, but, like *Gallionella*, it seems to thrive in environments that provide a gradient between reduced and oxidized forms of iron and manganese. *Leptothrix*

is also not a thermophilic form, but it has been observed in the recently discovered thermal springs on Loihi Seamount, the most eastward extension of Hawaiian-chain volcanism¹⁶. There, thick mats similar to the Crater Lake communities surround vents with temperatures that range between 15 and 30 °C (ref. 17).

In Table 1, the bulk composition of a bacterial mat is compared with that of hydrothermal iron oxide precipitates from several marine hydrothermal deposits and some iron-rich crusts collected in the vicinity of the Crater Lake bacterial mats. The mat contains >35% iron, which accounts for >70% of the mass of the mat (as FeOOH). Significant amounts of silica and phosphorus are also present. The relatively low sulphur content indicates that sulphur-oxidizing bacteria, which would transform hydrogen sulphide into elemental sulphur or sulphate, are not important members of the bacterial community. The concentrations of iron, aluminium and silicon in the mats and the hydrothermal oxide deposits are similar. The composition of the lithified crust is very similar to that of the bacterial mat.

Although the scanning electron micrographs demonstrate that the bacterial mats are predominantly composed of sheath-forming bacteria, only 2.4% of the material is organic carbon. It is probable that most of the bacteria shown in Fig. 2 are remnant sheaths rather than viable bacteria. Analyses of the organic fraction demonstrate that the extractable lipids are simple fatty acids. Sterols and other indicators of phytoplankton, zooplankton or local terrigenous sources are absent. In addition, the atomic ratio of carbon to nitrogen in the mat (2.9) is low compared with organic matter from plankton and is similar to bacterial sources of organic matter¹⁸. These analyses are consistent with fresh organic matter of microbial origin.

Within the general area where bacterial mats were found, we also observed lithified crusts of iron- and manganese-enriched material similar to the crust reported in Table 1. At some locations these deposits formed continuous pavements several metres across. At other sites multicoloured, pebble-sized material overlay what seemed to be normal sediments. The pebbles seem to be fragments of the lithified crusts.

Water associated with a bacterial mat was sampled by placing the open end of a cylindrical water sampler directly within the

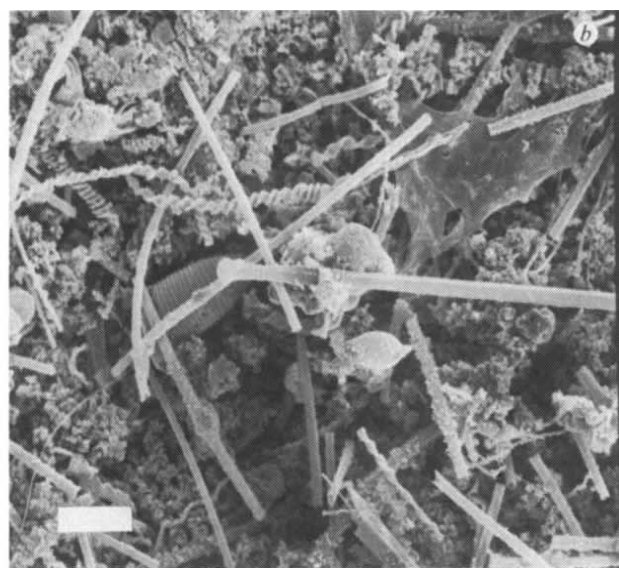
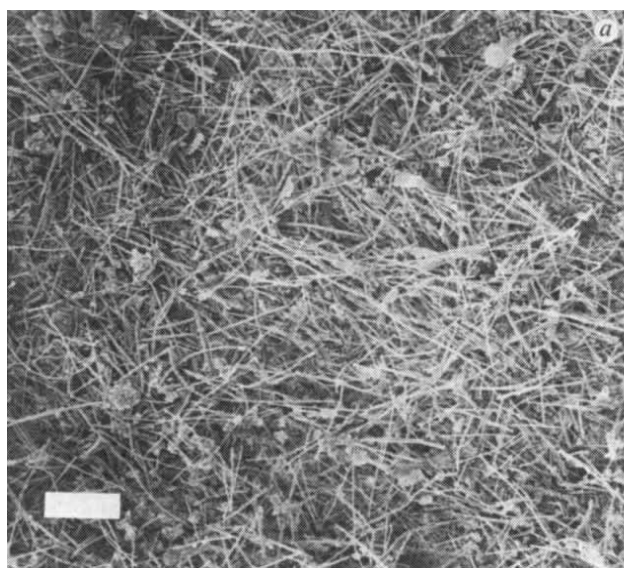


FIG. 2 a, A low-power scanning electron micrograph of the bacterial mat from dive 179. Tentative identification indicates these forms are iron-oxidizing sheath-forming bacteria of the genera, *Gallionella* and *Leptothrix*. The scale bar is 100 µm in length. b, A scanning electron micrograph of the bacterial mat from dive 179. The scale bar is 10 µm in length. The

genus, *Gallionella* are the forms with twisted sheaths; *Leptothrix* are the straight sheaths. The ribbed dome-like form in the centre is thought to be a protozoa. This form is seen on many of the scanning electron micrographs and could be feeding on the bacteria. The tube-like form in the centre-right with a blunt end is part of a diatom.

TABLE 2 Chemical composition of bacterial mat water and normal deep lake water

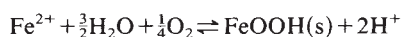
Sample	Na (mg l ⁻¹)	K (mg l ⁻¹)	Ca (mg l ⁻¹)	SO ₄ (mg l ⁻¹)	Cl (mg l ⁻¹)	SiO ₂ (mg l ⁻¹)	Li (μg l ⁻¹)	Mn (nM)
7.3 (hydrocast, 1987, 351 m)	10.4	1.71	6.73	10.2	10.6	17.3	44.1	0.6
CD179BT (bacterial mat water)	19.3	3.01	12.6	20.4	15.7	28.7	77.2	590

All analyses are by atomic absorption spectrophotometry.

mat (~10% of the sampler body penetrated the mat). After closing, the sampler retained several centimetres of mat material and a water sample that represents a mixture of the 'mat water' and of the local bottom water around the mat.

The bacterial-mat water sample contains the highest concentration of dissolved ions of any water sample collected within the caldera (Table 2). The concentrations of the major ions are approximately double the lake-water values, and the Mn concentrations are one thousand times higher than deep lake water. The mat water is an appropriate chemical 'endmember' that could explain the anomalously high concentrations of ions in Crater Lake¹⁰. Because of the sampling method, the extent of dilution of the bacterial-mat fluids by normal lake water is uncertain. A conservative estimate, based on relative volumes alone, suggests a 1:10 dilution of mat water by lake water. Thus, *in situ* mat-water concentrations are probably an order of magnitude larger than those shown in Table 2.

Elevated temperatures of fluids within the mat indicate geothermal heating. The contribution of heat by bacterial metabolic activity can be shown to be insignificant. (Assuming that the heat produced by oxidation of all dissolved ferrous ions in mat fluids is retained within the mat an upper limit to the metabolic heat production can be calculated. Using this closed-system assumption, the increase in temperature is a function of the concentration of iron in the fluids. An appropriate reaction is:



The enthalpy of this reaction¹⁹ is $-9.2 \text{ kcal mol}^{-1}$. Assuming an upper limit of 1.0 mM iron²⁰, conversion of all the dissolved iron to solid form would add ~9 cal of heat to each litre of water. Thus, the reaction would raise the temperature of that water by less than 0.01°. Chemical geothermometers^{21,22} permit an estimate of the water-rock reaction temperature of mat fluids. Assuming the mat-water sample (Table 2) was diluted ten times with normal lake water, the original concentration would be: Na = 99 mg l⁻¹; Ca = 65 mg l⁻¹; Mg = 37 mg l⁻¹; K = 15 mg l⁻¹; SiO₂ = 128 mg l⁻¹. If chalcedony (a microcrystalline form of quartz) is the phase controlling the silica content of the fluids, we estimate a temperature of 81 °C for the rock equilibration temperature of the mat fluid. Using the Na-K-Ca geothermometer^{21,22} (with no Mg correction) a temperature of 86 °C is calculated. There may be significant errors in the estimated temperatures because the assumed solubility and exchange reactions are slow to reach equilibrium at these relatively low temperatures and our endmember fluid composition is not well constrained. The result, however, is consistent with a geothermal source of the mat waters, and ~80 °C is our best estimate for the rock equilibration temperature of the geothermal fluids.

We conclude that the mats are produced by bacterial communities that oxidize ferrous iron carried in the fluids advecting through the mat. We base our assertion of fluid advection on the thermal and chemical gradients across the mats and the likelihood that the abundant growth of organisms requires a steady supply of reduced iron. We did not, however, measure or see evidence of flow through the mat, but dispersed and subtle flow of fluids would not be observable in the presence of the physical disturbance around the submarine.

We hypothesize that bacterial mats, lithified crusts and pebbled surfaces represent sequential stages in the evolution of

precipitates that form from heated fluids that advect through sediments and rocky outcrops at certain locations on the lake floor. The elevated temperatures and enrichment of these waters in elements commonly associated with thermal waters such as silicon, chlorine and lithium (Table 2) indicates the fluids are the result of hydrothermal circulation. Moreover, ³He, a definitive tracer of a mantle source, is strongly enriched in lake bottom waters near the bacterial mat⁹, compatible with hydrothermal input.

The following scenario can account for the presence of mats and crusts: iron-oxidizing, chemolithotropic bacteria precipitate large amounts of iron in mats at interfaces between the anoxic advecting fluids and oxic lake waters. Because these organisms require a source of ferrous iron, they thrive only as long as the advection of fluids continues. If the thermal-fluid advection is through sediments, as opposed to rock, excess pore pressure and doming of the sediments will result, similar to that observed in sediment-hosted hydrothermal systems in the oceans²³⁻²⁵. Precipitation within the conduits of hydrothermal systems, however, eventually blocks the flow and venting stops at a given site. When this occurs, the bacterial mat is transformed from a gelatinous living structure to a lithified crust. Cessation of venting would result in dome collapse and breaking of the lithified crusts would account for the pebble fields of mixed crust varieties. □

Received 28 July; accepted 27 October 1989.

- Bacon, C. & Lanphere, M. A. in *Crater Lake: An Ecosystem Study* (eds Drake, E. T., Larson, G. L., Dymond, J. & Collier, R.) (Am. Ass. Adv. Sci., San Francisco, in the press).
- Van Denburgh, A. S. *Geol. Surv. Prof. Pap. U.S.* No. 1005, 58-60 (1968).
- Simpson, H. J. thesis, Columbia Univ. (1970).
- Williams, D. L. & Von Herzen, R. P. *J. geophys. Res.* **88**, 1094-1104 (1983).
- Thompson, J. M., Nathenson, M. & White, L. D. in *Crater Lake: An Ecosystem Study* (eds Drake, E. T., Larson, G. L., Dymond, J. & Collier, R.) (Am. Ass. Adv. Sci., San Francisco, in the press).
- Nathenson, M. in *Crater Lake: An Ecosystem Study* (eds Drake, E. T., Larson, G. L., Dymond, J. & Collier, R.) (Am. Ass. Adv. Sci., San Francisco, in the press).
- Collier, R. W., Dymond, J., McManus, J. & Lupton, J. in *Crater Lake: An Ecosystem Study* (eds Drake, E. T., Larson, G. L., Dymond, J. & Collier, R.) (Am. Ass. Adv. Sci., San Francisco, in the press).
- Collier, R. *et al. Eos* **68**, 1721 (1987).
- Lupton, J. E., Collier, R. & Dymond, J. *Eos* **68**, 1722 (1987).
- Collier, R. W. & Dymond, J. *Studies of Hydrothermal Processes in Crater Lake Rep. No. 79 CPSU* (Oregon State Univ., Corvallis, 1989).
- Ehrlich, H. L. *Geomicrobiology* 393 (Dekker, New York, 1981).
- Hanert, H. H. in *The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria* (eds Starr, M. P., Stolp, H., Truper, H. G., Balows, A. & Schlegel, H. G.) 509-515 (Springer, Berlin, 1981).
- Wolfe, R. S. in *Principles and Applications in Aquatic Microbiology* (eds Heukelakian, H. & Dondero, N. C.) 82-97 (Wiley, New York, 1964).
- Hanert, H. H. *Geol. Rdsch.* **62**, 786-812 (1973).
- Mulder, E. G. & Deinema, M. H. in *The Prokaryotes: A Handbook on Habitats, Isolation, and Identification of Bacteria* (eds Starr, M. P., Stolp, H., Truper, H. G., Balows, A. & Schlegel, H. G.) 425-440 (Springer, Berlin, 1981).
- Karl, D. M., Brittain, A. & Tilbrook, B. *Global Venting, Midwater and Benthic Ecological Processes* (eds DeLuca, M. P. & Babb, I.) (U.S. Dept of Commerce, Washington, DC, 1988).
- Karl, D. M., McMurtry, G. M., Malahoff, A. & Garcia, M. O. *Nature* **335**, 532-535 (1988).
- Goldman, J. C., Caron, D. A. & Dennett, M. R. *Limnol. Oceanogr.* **32**, 1239-1252 (1987).
- Garrels, R. M. & MacKenzie, F. T. in *Equilibrium Concepts in Natural Water Systems (151st Meet. Am. chem. Soc.)* 222-242 (Am. chem. Soc., Washington, DC, 1967).
- Mariner, R. H., Presser, T. S. & Evans, W. C. *Geol. Surv. Open-File Rep.* 82-98 (1982).
- Fournier, R. O. in *Geothermal Systems: Principles and Case Histories* (eds Ryback, L. & Muffler, L. P. J.) (Wiley, New York, 1981).
- Fournier, R. O. & Potter, R. W. *Geochim. cosmochim. Acta* **43**, 1543-1550 (1979).
- Williams, D. L. *et al. J. geophys. Res.* **84**, 7467-7484 (1979).
- Maris, C. R. P., Bender, M. L., Froelich, P. N., Barnes, P. N. & Luedtke, N. A. *Geochim. cosmochim. Acta* **48**, 2331-2346 (1984).
- Wheat, C. G. & McDuff, R. E. *Eos* **69**, 1499 (1988).
- Corliss, J. B., Lyle, M., Dymond, J. & Crane, K. *Earth planet. Sci. Lett.* **40**, 12-24 (1978).
- Bostrom, K. & Widenfalk, L. *Chem. Geol.* **42**, 203-218 (1984).

ACKNOWLEDGEMENTS. We thank the US National Park Service employees for support, R. Conard, J. McManus and C. Meredith for chemical analyses, C. Dahm for help with bacterial identification and a review, and D. Karl, S. Roth and R. McDuff for reviews. This study was funded by the US National Park Service.

The earliest known reptile

T. R. Smithson

Cambridge Regional College, Newmarket Road, Cambridge CB5 8EG, UK

AMNIOTES (reptiles, birds and mammals) are distinguished from non-amniote tetrapods (amphibians) by the presence of complex embryonic membranes. One of these, the amnion, gives its name to the group. Very few skeletal characters distinguish amniotes from amphibians¹, making it difficult to recognize early amniotes in the fossil record. The earliest amniote fossil identified so far is *Hylonomus* from the Westphalian (Upper Carboniferous) of Joggins, Nova Scotia^{2,3}, (~300 Myr). I report here the discovery of a much earlier amniote skeleton from the Brigantian (Lower Carboniferous) of Scotland (~338 Myr)⁴, which thus represents the earliest occurrence of amniotes in the fossil record. The specimen was collected from the East Kirkton Limestone, near Bathgate, West Lothian⁴⁻⁸, and is part of a unique terrestrial fauna that includes eurypterids, myriapods, scorpions and the earliest-known harvestman spider^{7,9}, together with the earliest known temnospondyls, a group that may include the ancestors of all living amphibians¹⁰. It will make an important contribution to our knowledge of early amniote morphology and the interrelationships of tetrapods.

The specimen described here (SPW2326) currently forms part of the private collection of Stanley Wood who found it in the black shale member of the East Kirkton succession. It is preserved in part and counterpart and consists of an almost entire articulated skeleton (Fig. 1) which has been split into halves so that little surface detail of the bones is visible. The bedding plane in which the specimen lay is slightly contorted, disrupting the anterior parts of the skeleton. The total length of the skeleton is between 180–200 mm (presacral length ~120 mm).

The skull has been crushed dorsoventrally. Most of the skull roof is preserved, together with the right cheek and mandible which have been displaced laterally (Fig. 2a). Little of the palate is visible, but there is evidence that it is concealed beneath the overlying bones of the skull roof. The snout–occiput length is ~18 mm and the snout–quadrate length ~20 mm. The pre-orbital region shows a pattern of bones common to all early tetrapods, with midline frontals and nasals flanked laterally by pre- and postfrontals and large premaxillae anteriorly. But,

unlike other Carboniferous amniotes such as protorothyridids¹¹, the pre- and postfrontals are in contact, preventing the exposure of the frontal in the orbital margin (Fig. 2a). In the post-orbital region the midline parietals appear to contact the postorbital and squamosal. There is no evidence that the temporal series (intertemporal, supratemporal and tabular) intervenes between the parietals and cheek bones. The squamosal, postorbital, jugal and quadratojugal are arranged in the primitive tetrapod pattern¹². Nine sclerotic plates are preserved in the dorsal half of the orbit.

The vertebral column is preserved in articulation, although in places the bone is eroded, making it difficult to distinguish individual bones. There are approximately 32 presacral vertebrae and a similar number make up the preserved portion of the tail. The first eight presacral vertebrae are preserved in dorsal view, the remainder in lateral aspect. The form of the vertebrae is illustrated in Fig. 2b. The hourglass-shaped centrum seems to be fused to the neural arch which bears a short neural spine. Between each successive centrum is a wedge-shaped intercentrum. This is relatively much larger than those of the Westphalian protorothyridids², but much smaller than those of other reptiliomorphs from East Kirkton^{6,7}. The neural arches are narrow and unswollen, and all the presacral vertebrae seem to bear ribs.

The pectoral girdle and fore limbs are incompletely preserved, and will require detailed study before they can be adequately described, but the pelvic girdle and hind limbs are very well preserved. They lie below the vertebral column except that the left pes overlies vertebrae in the sacral region, obscuring detail of the pattern of its bones. The right pelvis and hind limbs are illustrated in Fig. 2c. The camera lucida drawing was made of the bones as preserved on the counterpart block and thus appears to be from the left side of the skeleton. The dorsoventrally flattened pelvis consists of a large puboischiadic plate 16 mm long and a relatively short, narrow iliac blade ~6 mm long. The limb bones are long and gracile. The femur is 11 mm long and the tibia and fibula are each ~6 mm long. The right pes is very well ossified and the majority of bones are preserved. Two large proximal tarsal bones are readily identified and are almost certainly the homologues of the astragalus and calcaneum of extant amniotes. Most of the centralia are preserved in close apposition to the metatarsals, but the distal bones of the foot have folded back towards the epipodials. The digits are very

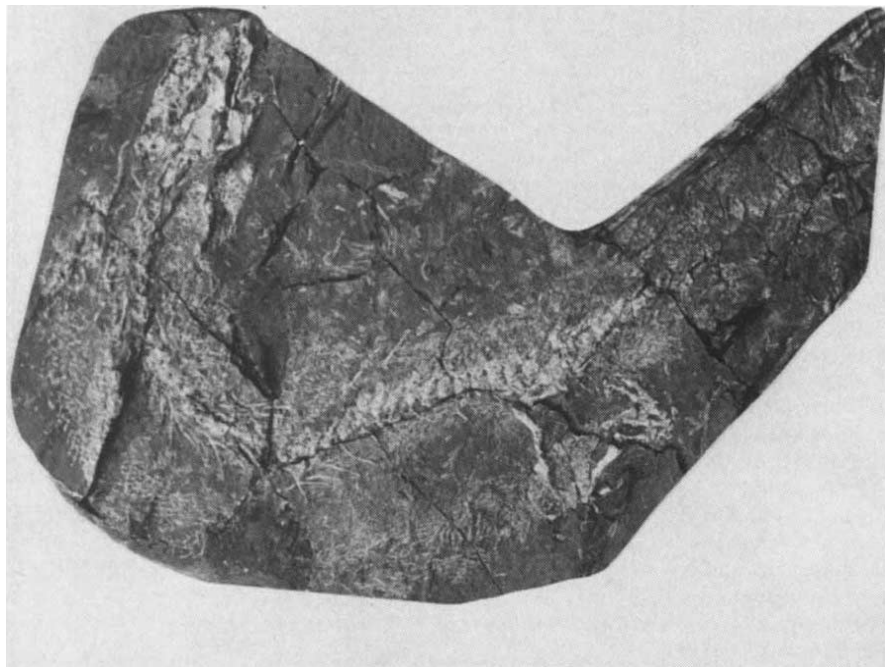


FIG. 1 Fossil amniote skeleton from the East Kirkton Limestone (SPW2326). Counterpart block. Scale bar, 50 mm.

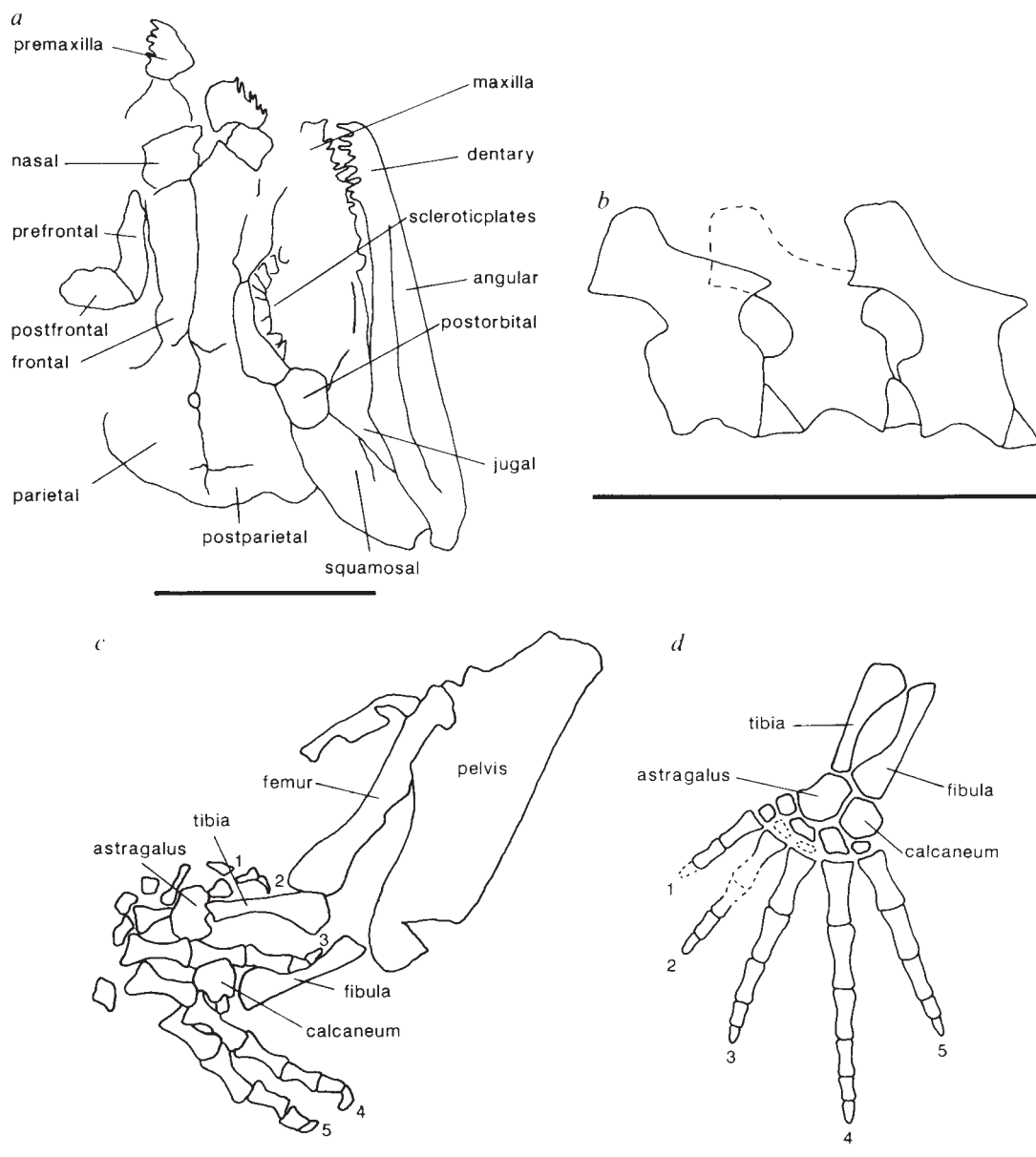


FIG. 2 Fossil amniote (SPW2326). *a*, Skull, dorsal view. Composite figure based on data from part and counterpart halves. *b*, Trunk vertebrae (approximate presacral numbers 20–22), lateral view. *c*, Right pelvis and hind limb drawn from counterpart block. *d*, Restoration of right tibia, fibula and pes, reversed. Scale bars, 10 mm.

well-articulated and the pattern of phalanges is clear in all but the first digit, in which the terminal phalanx appears to be missing. The phalangeal formula is 23454. A restoration of the tibia, fibula and pes is shown in Fig. 2*d*.

SPW2326 is placed in the Amniota on the basis of the presence of an astragalus and calcaneum (a character it shares with extant amniotes) and the reduction of the temporal series to permit contact between the parietal, the postorbital and the squamosal (a character shared with other Carboniferous amniotes). Among Palaeozoic tetrapods, the presence of an astragalus and calcaneum is not a unique feature of amniotes: these bones have been described in a number of microsaur¹³. But the microsauro skull is characterized by a single large bone in the temporal series, usually described as the tabular, which separates the midline parietal from the postorbital and squamosal bones of the cheek. In addition, no microsauro has been described with a pedal phalangeal formula 23454, although this pattern has been restored in *Tuditanus*¹⁴. No other features are evident on SPW2326 that would indicate relationship with other groups of Carboniferous amphibians. The pattern of the skull roofing bones of SPW2326 precludes relationship with one of the groups of non-amniote reptiliomorphs (see ref. 1) or with the taxon including ichthyostegids, colosteids and temnospondyls. It lacks

the unique vertebral morphology of nectrideans¹⁵ and the aberrant cranial morphology of lysorophids and adelogyrids¹³, whereas the presence of limbs excludes it from the Aistopoda.

SPW2326 is thus considered to be an amniote. However, its position within the Amniota remains to be determined and must await detailed description. The absence of lateral temporal openings in the skull clearly precludes inclusion in the Synapsida or Diapsida, but its relationships with the other two major groups of Carboniferous amniotes, the Captorhinidae and Protorothyrididae¹¹ are unclear. It remains to be seen whether the discovery of SPW2326 will help to resolve the problem of the relationship of amniotes to non-amniotes. Recent discussion of this problem has been dominated by an analysis of characters from the braincase and otic region of Upper Carboniferous and Lower Permian taxa^{1,16,17}. But, as the whole skeleton of the earliest amniote becomes better known, and the other reptiliomorphs from East Kirkton are described, it is anticipated that new characters will be discovered which will provide tests for current phylogenetic hypotheses. □

Received 21 August; accepted 27 October 1989.

1. Panchen, A. L. & Smithson, T. R. in *The Phylogeny and Classification of the Tetrapods*, Vol. 1: *Amphibians, Reptiles, Birds* (ed. Benton, M. J.) 1–32 (Clarendon, Oxford, 1988).
2. Carroll, R. L. *J. Linn. Soc. (Zool.)* **45**, 61–83 (1964).

3. Dawson, J. W. *Air-breathers of the Coal Period* (Dawson Brothers, Montreal, 1863).
4. Wood, S. P., Panchen, A. L. & Smithson, T. R. *Nature* **314**, 355–356 (1985).
5. Smithson, T. R. *Scott. J. Geol.* **21**, 123–142 (1985).
6. Milner, A. R. *et al. Mod. Geol.* **10**, 1–28 (1986).
7. Rolfe, W. D. I. *Earth Sci. Cons.* **25**, 22–28 (1988).
8. Muir, R. O. & Walton, E. K. *Trans. geol. Soc. Glasg.* **22**, 157–168 (1957).
9. Rolfe, W. D. I. *et al. in Volcanism and Fossil Biotas* (ed. Lockley, M.) (Geological Society of America, Special Paper 244, in the press).
10. Milner, A. L. in *The Phylogeny and Classification of the Tetrapods, Vol. 1: Amphibians, Reptiles, Birds* (ed. Benton, M. J.) 59–102 (Clarendon, Oxford, 1988).
11. Heaton, M. J. & Reisz, R. R. *Can. J. Earth Sci.* **23**, 402–418 (1986).
12. Panchen, A. L. & Smithson, T. R. *Biol. Rev.* **62**, 341–438 (1987).
13. Carroll, R. L. & Gaskill, P. *Mem. Am. phil. Soc.* **126**, 1–211 (1978).
14. Carroll, R. L. & Bard, D. *Amer. Mus. Novit.* **2337**, 1–50 (1968).
15. Milner, A. C. in *The Terrestrial Environment and the Origin of Land Vertebrates* (ed. Panchen, A. L.) 377–405 (Academic, London, 1980).
16. Heaton, M. J. in *The Terrestrial Environment and the Origin of Land Vertebrates* (ed. Panchen, A. L.) 491–551 (Academic, London, 1980).
17. Smithson, T. R. *Zool. J. Linn. Soc.* **85**, 317–410 (1985).

ACKNOWLEDGEMENTS. I am grateful to Mr Stanley Wood for the invitation to write this preliminary description of his new specimen and Dr Ian Rolfe for providing facilities at the Royal Museum of Scotland during its study. I thank Drs Robert Carroll, Andrew Milner, Alec Panchen and Mr Wood for helpful criticism of earlier drafts of the manuscript.

Evidence from aphasia for the role of proper names as pure referring expressions

Carlo Semenza & Marina Zettin*

Dipartimento di Psicologia Generale, Piazza Capitanato,
3, 35139 Padova, Italy

HERE we describe a patient who, as a result of brain damage, had a dramatic inability to retrieve proper names, and who thus offered the opportunity of observing the distinction made in the brain between proper and common names. Although category specific aphasic disturbances are relatively common^{1,2}, this patient's anomia for proper names is very rare: only two other cases have been described so far^{3,4}. The opposite phenomenon, a selective sparing of proper names, has also been recently observed⁵. It is widely assumed that double dissociations such as this reflect the premorbid organization of the cognitive system: the categories of proper and common nouns, therefore, would be separately represented or, at least, separately accessed in the intact brain. The only other such deficits, including the inability to learn arbitrary links between words and to guess the titles of pieces of music, are believed to be indicative of a problem in dealing with purely referential relations. This, in turn, would indirectly confirm the role of proper names as pure referring expressions.

L.S. was a 41-year-old employee in a hardware store, with eight years of education, who suffered a head trauma after a riding accident in July 1988. At that time a computerized tomography documented a left fronto-temporal lesion; in September 1989 NMR scanning showed a similar picture, emphasizing the involvement of basal structures. Clinical examination revealed no neurological signs, and no psychiatric pathological signs were ever observed. Very mild word-finding difficulties were noted. In September 1988 no pathological signs were detectable in his spontaneous speech or in his language comprehension. His residual deficit consisted almost exclusively of a dramatic inability to retrieve proper names either spontaneously or on formal testing. This deficit affected only oral and written word retrieval, because his auditory and written comprehension of proper names were both intact. On neuropsychological examination with standard test batteries^{6,7} he obtained normal scores: in particular his visual recognition (including faces), spatial imagery, short- and long-term memory (with the exceptions reported below) were found to be intact. He accomplished perfectly all standard linguistic tasks. The following is a descrip-

tion of his inability to retrieve proper names and of his associated disorders.

In visual-confrontation naming he was 100% correct with real objects (50 items) and he scored 98 out of 100 with pictorial material. Categories shown in previous studies to be often selectively sensitive to brain damage¹ were also separately tested in a confrontation naming paradigm. A minimum of ten items for each of these categories were used as stimuli: L.S. made no errors. His performance with proper names was completely different. With real persons his knowledge of names was limited to those of most members of his family, which he had relearned during the months before testing. He gave correct information, however, about other real persons ("doctor"), and never experienced recognition problems. When presented with pictures of famous persons his naming performance was 2/25. Nonetheless, he provided correct information and details about these people's identities in every case (for example "prime minister"). In all cases he had no problems in matching spoken-aloud or written names with the pictures in a multiple-choice test. This ability to retrieve geographical names was tested using a blank atlas and pictures of the sites of well-known cities. He could name only 6/15 pictures and 4/15 map sites, but he was able to match them correctly in pairs. He also gave correct information about all of the items. In addition he was tested using the blank atlas on rivers (naming 3/8), countries (2/8) and mountains (2/8). Again, for all items, he provided a vast amount of correct information. We confronted him with several popular pieces of wordless classical music. If he demonstrated that he recognized them by singing along and continuing the melody appropriately when the music stopped, then we asked him to name the piece. He was unable to retrieve the title of any of such 12 pieces, regardless of whether their proper names were very abstract ('5th Symphony') or descriptive ('Moonlight Sonata'). He was also unable to provide the names of the composers.

In naming from definition L.S.'s performance was again very satisfactory provided that proper names were not the target. His score was 75/80, the stimuli being various categories of nouns (including animate beings, inanimate objects and abstract nouns), adjectives, verbs and numbers. By contrast, his performance with proper names was very poor; he scored 5/8 for relatives, 2/15 for famous persons, 3/5 for cities, 1/5 for mountains and 3/5 for countries. He could point correctly to all the geographical locations on the blank atlas. Phonological cueing (consisting of the first phoneme of the name) alone never improved his performance in generating proper names on either confrontation or definition. Semantic cueing was more effective, but only in the particular case in which the proper name also had a real meaning. For instance, Colombo (Columbus) in Italian means 'pigeon': L.S. seemed to be helped if the description included a cue to the meaning of the name ("Tell me who discovered America. He had the name of a bird.") His performance was correct in 4 out of 15 items of this kind, and when helped also with a phonetic cue he was able to score 9/15.

Naming by category in one minute was also tested. In this case he provided the names of 12 fruits, 11 vegetables, 9 birds, 13 clothes, 6 forms of footwear, 16 body parts, 15 means of transportation, 13 occupations, 12 sports, 7 carpenter's tools, 10 types of pasta, 13 common names beginning with the letter 'F', 9 beginning with 'M' and 10 beginning with 'S'. In the same time limit he was able to generate 12 proper names when "any person's name" was requested, but his performance dropped dramatically when more delimited categories were asked (one musician, two sportsmen, two actors, and two politicians). Additional time did not improve his performance with proper name categories, whereas for all other categories he continued to give new correct items.

L.S. was able to repeat immediately any proper name and to retrieve it if allowed to rehearse it silently without interference. Any interference procedure (such as counting backwards),

* Present address: Clinica Ausiliatrice, Via Peyron, Torino, Italy.

however, caused his performance to drop to zero in 15 seconds.

The ability to acquire arbitrary associations between pairs of words, which may be considered as conceptually equivalent to naming proper names, was finally tested. The patient was given the paired-associate learning task from the Wechsler Memory Scale and the Bisiach, Cappa and Vallar⁷ battery. He quickly learned the 'easy' pairs (in which the two words have a semantic relation to each other, such as 'north'-'south'), but he showed a chance performance over several attempts in retrieving the second member of the pair, given the first, when the relation between the two was totally arbitrary (such as 'clock'-'pear'). By contrast, he had no difficulty in a word-learning task. Specifically his score on a ten-word free-recall task (three trials) was within the normal range⁷. At work he also had great difficulty in learning numerical labels consistently when applied to certain items of hardware.

Our evidence indicates a common denominator underlying L.S.'s performances: the inability to deal (at the retrieval level) with purely referential nondescriptive semantic relations. Indeed, proper names are believed by some philosophers of language^{8,9} to be pure referring expressions. According to Frege's¹⁰ distinction, proper names just carry 'reference', that is they denote the individuals or the entities that are called by them, but have no 'sense', that is, they do not describe any property or imply any attribute. They are the opposite of 'descriptions', which have sense and which encompass all common nouns. Meaningless labels, which is what pure referring expressions are, could indeed need an independent system to take care of them; in L.S. this system would be disrupted. Psychologically, this makes more sense than the alternative explanation that L.S.'s inabilities arose because of a difficulty purely in retrieving proper names; in this case one would be left with the problem of specifying what particular quality makes it useful for the brain to implement proper names independently from common ones. The need of an independent system may still be there even if proper names were not meaningless labels but 'shorthand descriptions', as some authors¹¹ prefer to the pure reference theory. A descriptive value, varying from none (or very little) to presumably a lot, would be in any case an important dimension in the organization of semantic memory. It would be orthogonal with other dimensions such as the one defined by the prevalence of structural or functional attributes, which also has been suggested to determine category effects in aphasia¹².

There are at present only a few highly speculative models of the actual wiring in the nervous system of such organization: those involving modularly distinct subsystems^{12,13} or regions of relative specialization within a distributed net^{2,5} are the more accredited ones.

Note added in proof. Note also that a substantially similar argument comes from modern linguistics. In Jackendoff's theory of natural language semantics, an important distinction in conceptual structure is the binary feature Type/Token¹⁴. What one learns and stores in memory can be linked either with Token (if one is remembering an individual) or with Type (if one has learned a category). Proper names appear typically to denote individuals. □

12. Warrington, E. K. & Shallice, T. *Brain* **107**, 829-853 (1984).
13. Warrington, E. K. & McCarthy, R. A. *Brain* **106**, 859-878 (1983).
14. Jackendoff, R. *Semantic and Cognition* (MIT Press, Cambridge, Massachusetts, 1983).

ACKNOWLEDGEMENTS. We are grateful to Professor Roberto Rago of the Clinica Ausiliatrice, Torino, for providing facilities to carry out this study.

Analysis of fusion gene and encoded photopigment of colour-blind humans

Jay Neitz, Maureen Neitz & Gerald H. Jacobs

Department of Psychology and Neuroscience Research Institute, University of California, Santa Barbara, California 93106, USA

IN humans, long-wavelength-sensitive and middle-wavelength-sensitive cone pigments are encoded by genes lying in a head-to-tail tandem array on the X chromosome. Deficiencies in red-green colour vision seem to arise from unequal recombination of these normal X-linked genes^{1,2}. In some dichromats this recombination is believed to yield a fusion gene encoding a product with an absorption spectrum similar to that of one or the other of the normal photopigments². Until now, however, such a relationship between the structure of a pigment gene and the spectral properties of its encoded pigment has not been directly shown. We have now sequenced a fusion gene isolated from a red-green colour-blind human and determined the spectral properties of the pigment that it encodes. The absorption spectrum of the photopigment was very similar to that of normal middle-wavelength-sensitive photopigment, even though about half of its DNA coding sequence seems to be derived from a gene encoding normal long-wavelength-sensitive pigment. These results indicate the regions of the X-encoded photopigment apoproteins that are responsible for differences in their spectral tuning, and imply that the striking variations in colour vision among anomalous trichromats of a particular type are not attributable to anomalous pigments with differing spectral peaks.

We established the colour vision of a 31-year-old male by psychophysical tests in which the appearances of mixtures of red and green primary lights were compared with the appearance of a standard yellow light (the Rayleigh match). He accepted

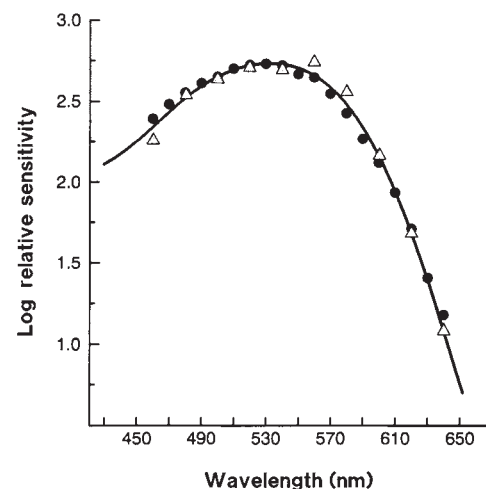


FIG. 1 Comparison of the spectral sensitivity of the MWS cone pigment from a protanope with the MWS cone pigment from a colour-normal human. ●, Protanope ERG sensitivity values corrected for pre-retinal absorption by the lens¹². Δ, Spectral sensitivity of MWS cone pigment from a normal trichromat⁵. The wavelength of peak sensitivity for each was determined by translating a standard visual pigment absorption curve¹³ on a log-wavenumber axis to obtain the best fit. The two pigment spectra have the same best-fitting curve (solid line; $\lambda_{max} = 530$ nm).

Received 4 July; accepted 17 October 1989.

1. Semenza, C. & Denes, G. *Aphasiology* **2**, 405-410 (1988).
2. McCarthy, R. A. & Warrington, E. K. *Nature* **334**, 428-430 (1988).
3. McKenna, P. & Warrington, E. K. *J. Neurol. Neurosurg. Psychiatr.* **43**, 781-788 (1978).
4. Semenza, C. & Zettin, M. *Cognitive Neuropsychol.* **5**, 711-721 (1988).
5. Warrington, E. K. & McCarthy, R. A. *Brain* **110**, 1273-1296 (1987).
6. Goodglass, H. & Kaplan, E. *The Assessment of Aphasia and Related Disorders*. (Lea & Febiger, Philadelphia, 1972).
7. Bisiach, E., Cappa, S. & Vallar, G. *Guida all'esame neuropsicologico* (Cortina, Milan, 1983).
8. Mill, J. S. *A System of Logic* (Longman, London, 1843).
9. Kripke, S. *Naming and Necessity* (Blackwell, Oxford, 1980).
10. Frege, G. in G. Frege, *Funktion, Begriff, Bedeutung* (ed. Patzig, G.) 40-65 (Vandenhoeck und Ruprecht, Göttingen, 1982).
11. Searle, J. R. in *Semantics, an Interdisciplinary Reader in Philosophy, Linguistics and Psychology* (eds Steinberg, D. D. & Jakobovitz, L. A.) 134-141 (Cambridge University Press, 1971).

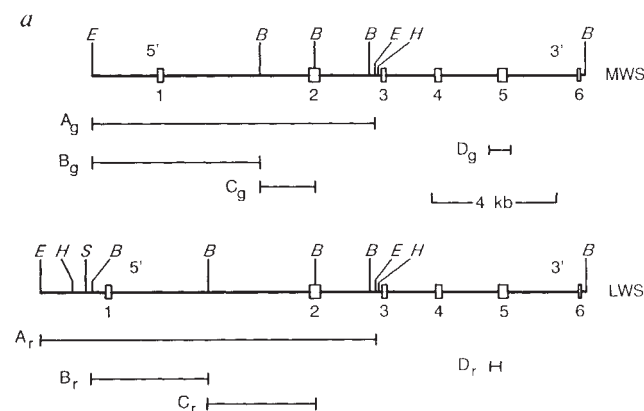


FIG. 2 a, Restriction maps and four restriction length polymorphisms that have been used to distinguish human X-linked cone pigment genes (from ref. 1). *B*, *Bam*HI; *E*, *Eco*RI; *H*, *Hind*III; *S*, *Sal*I. *b*, *c*, Two Southern blots of genomic DNA from the protanope (P) compared with identical digests of DNA from two colour-normal humans (N). *A_r* and *A_g* are *Eco*RI fragments visualized with a complementary DNA probe encompassing exon 1 and the 5' half of exon 2. *D_g* and *D_r* are *Rsa*I fragments visualized with a probe from the 3' end of intron 4.

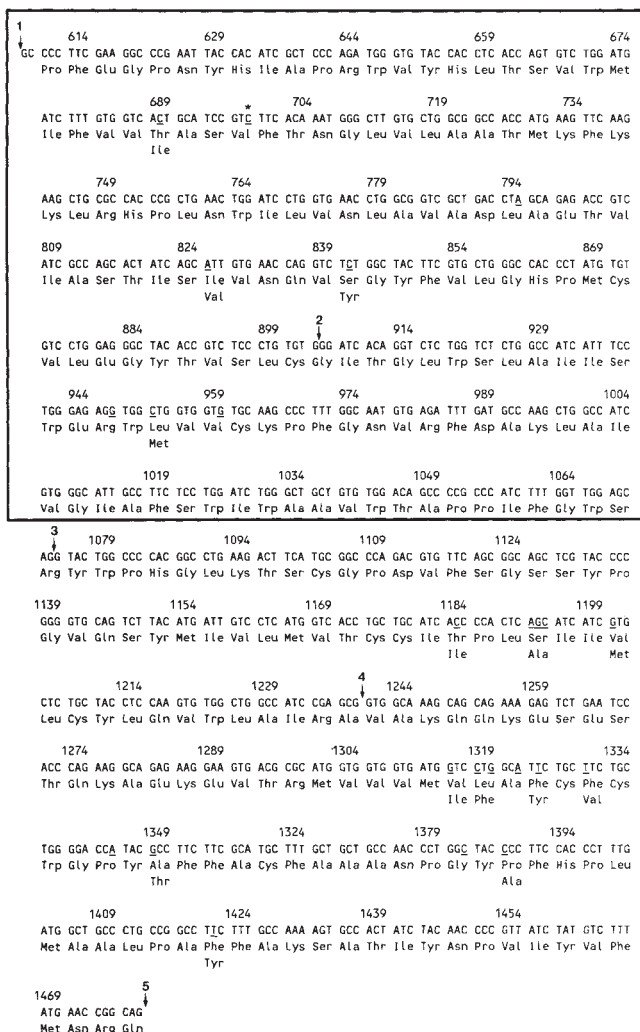
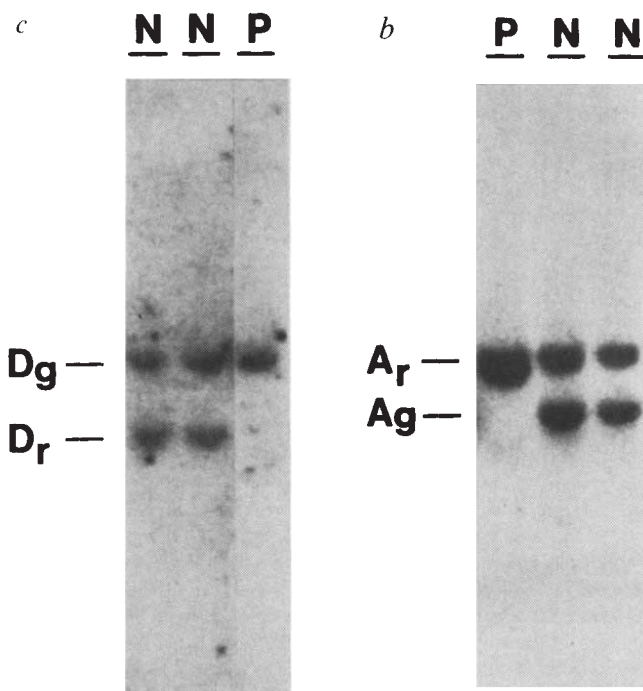


FIG. 3 Sequence of exons 2, 3, 4, and 5 of the MWS pigment gene from a protanope. The locations of the five introns¹ are indicated by arrows. The numbering convention is that of Nathans *et al.*¹. For the region outside the box (exons 4 and 5), the protanope gene sequence is identical to that of an identified normal MWS pigment gene fragment (gJHN21; ref. 1). For these two exons the underlined nucleotides in the protanope gene sequence are those that differ from corresponding nucleotides in a genomic LWS pigment gene fragment (gJHN33; ref. 1), and the lower amino-acid sequence is that of the identified normal LWS pigment. Exons 2 and 3 (enclosed in the box) of the protanope gene are identical to those of the identified LWS pigment genes in all of the nucleotide substitutions that distinguish the identified normal MWS from the LWS pigment genes¹. But exons 2 and 3 of the protanope gene differ slightly from those of each of the three genes (one derived from a genomic library and two derived from a human retina cDNA library, no two of which were identical) that Nathans *et al.*¹ identified as encoding normal LWS pigment. For example, exons 2 and 3 of the protanope gene differ from those of the genomic LWS pigment gene fragment (gJHN53; ref. 1) by three nucleotides, but for two of these substitutions the protanope gene is identical to the other two LWS pigment genes, whereas the third is a silent nucleotide substitution (indicated by the asterisk) in which the protanope gene sequence differs from both MWS and LWS sequences. For exons 2 and 3 the nucleotides of the protanope gene that differ from those of a MWS gene fragment (gJHN21; ref. 1) are underlined, and the lower amino-acid sequence is that deduced from the MWS pigment gene (identical amino-acid residues are left blank). A genomic library derived from germ line DNA from the protanope was constructed¹⁴ in the cosmid vector plasmid pWE15 (ref. 15). A fragment containing the protanope MWS pigment gene was identified by its hybridization with a cDNA clone probe. Sequences (both strands) were determined by the dideoxy chain termination method¹⁶.

the red and green mixture of all proportions as matching the yellow light in both small-field (2 degrees) and large-field (11 degrees) versions of the test. The subject was thus a true dichromat^{3,4}—a protanope, having only a single X-encoded cone pigment absorbing light of middle to long wavelengths.

We determined the absorption spectrum of this X-encoded cone pigment by analysing a retinal gross potential—the electroretinogram (ERG). To measure relative sensitivity, we adjusted the quantal intensity of a rapidly flickering monochromatic test light to produce a response that matched in amplitude the response to an interleaved flickering reference light⁵. Repetition of this procedure for many test wavelengths yields a curve that accurately characterizes absorption spectra of middle-wavelength-sensitive (MWS) and long-wavelength-sensitive (LWS) cone pigments in dichromatic subjects⁵. The derived sensitivity of the dichromat curve (Fig. 1, ●) thus gave the spectral sensitivity of this MWS pigment. Schnapf *et al.*⁶ have measured the spectral sensitivity of normal human cones. The absorption spectra for the protanope MWS cone and the normal MWS cone were very similar (Fig. 1) and the spectral peaks were identical ($\lambda_{\max} = 530$ nm).

Southern blots for two different restriction-enzyme digests of genomic DNA from the protanopic subject are shown in Fig. 2. For humans with normal colour vision, similar analysis reveals two hybridizing bands corresponding to two classes of X-linked pigment genes¹. For the protanopic subject, we observed only a single hybridizing band for each digest, indicating the presence of a single X-linked pigment gene. The earlier identification of the normal pigment genes^{1,2} (Fig. 2a) enabled us to identify this pigment gene as a fusion gene which had derived its 5' end from a normal LWS pigment gene and its 3' end from a normal MWS pigment gene.

We obtained a DNA fragment containing the entire X-linked photopigment gene of the protanope from a library constructed using total genomic DNA, and determined the sequences of exons 2, 3, 4 and 5. Comparison with the normal X-linked pigment genes¹ indicated that this fusion gene had derived its exons 2 and 3 from a normal LWS pigment gene and its exons 4 and 5 from a normal MWS pigment gene (Fig. 3). The fact that the absorption spectrum of the pigment encoded by the fusion gene does not differ from that of a normal MWS pigment indicates that the amino-acid substitutions encoded in exons 2 and 3 have little or no effect on the absorption spectrum of the X-linked photopigments. Nathans *et al.*¹ found no differences between exons 1 or between exons 6 of the MWS and LWS pigment genes. Thus the spectral differences between MWS and LWS pigments seem to be governed entirely by the amino-acid substitutions encoded in exons 4 and 5.

Except for several homologous substitutions⁷ (which are not expected to alter the absorption spectrum of the chromophore) all the amino-acid differences between MWS and LWS photopigments that occur in the transmembrane segments encoded by exons 4 and 5 are substitutions of nonpolar groups for hydroxyl-bearing groups¹. Studies of bacteriorhodopsin demonstrate that similar substitutions of nonpolar for polar groups (for example, Phe for Trp) can have dramatic effects on the absorption spectrum of the chromophore⁸. Depending on their position relative to the chromophore, hydroxyl groups have been proposed to preferentially stabilize or destabilize the ground state relative to the excited state and thus shift the absorption spectrum to the shorter or longer wavelengths⁹. Figure 4 shows the location of all the nucleotide differences responsible for non-homologous amino-acid substitutions that are located in the transmembrane segments encoded by exons 4 and 5. There are two such substitutions in exon 4 and three in exon 5. The most 5' proximal of these nucleotide substitutions in exon 5 is marked by the restriction site polymorphism that distinguishes fragment D_r from D_g .

The locations and small number of nucleotide substitutions encoding the amino acids that are candidates for producing the

spectral difference between X-linked pigments (Fig. 4) limits the number of different pigments with intermediate spectral peaks that can result from genetic recombination. These limitations have implications for understanding the genetic mechanisms underlying the most common forms of defective colour vision—the anomalous trichromacies. Nathans *et al.*² propose that the anomalous pigments result from recombination of the identified normal pigment genes. Because all protanomalous individuals that have been studied using Southern blot analysis have fusion genes containing fragment D_g , and all deuteranomalous subjects have fusion genes containing fragment D_r (refs 2, 10), both types of anomalous pigment genes must be produced by crossovers upstream of the D_g - D_r marker. The nucleotide substitutions that give rise to the two hydroxyl-bearing amino-acid substitutions in the MWS exon 4 are separated by only seven nucleotides, so recombination events in the region between them are expected to occur only rarely. These observations suggest that one or both of the substitutions encoded in exon 4 are responsible for the small spectral shift differentiating the X-linked pigments in anomalous trichromats, whereas one or more of the substitutions encoded in exon 5 are

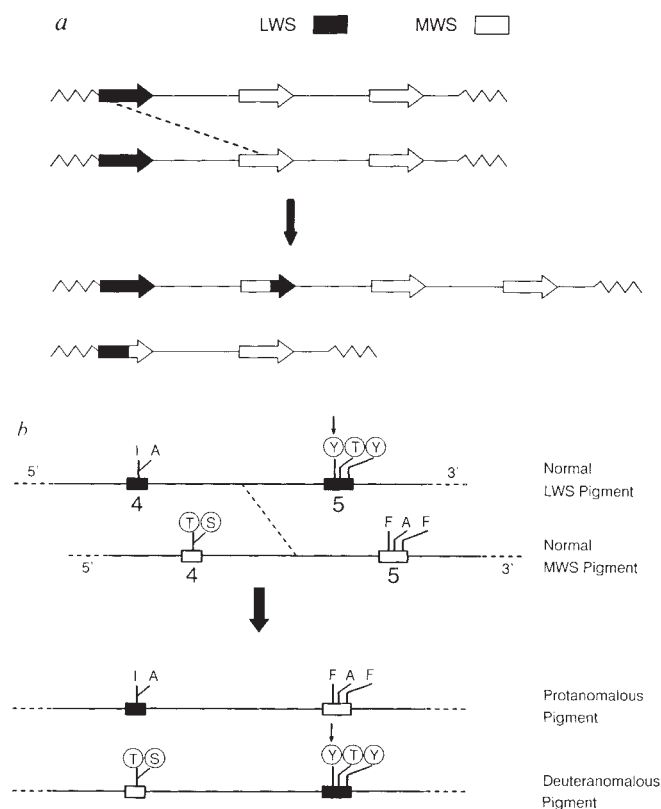


FIG. 4 a, Scheme showing how unequal intragenic recombination between normal MWS and LWS pigment genes can produce fusion genes. Each arrow represents a single gene (from ref. 2). b, Detail of a, showing only exons 4 and 5 of the pigment genes. The locations of nucleotide differences encoding amino acid substitutions that are candidates for producing the spectral shifts differentiating the products of X-linked pigment genes are labelled according to the amino acid produced. The hydroxyl-bearing amino acids are circled (Y, tyrosine; T, threonine; S, serine), and their nonpolar counterparts are not circled (I, isoleucine; A, alanine; F, phenylalanine). The location of the *Rsa*I site that differentiates restriction fragment D_r from D_g (Fig. 1) is indicated by a small arrow. Cross-overs upstream of the two substitutions in exon 4 produce fusion genes that encode pigments with spectra unchanged from those of products encoded by the original genes. Exchanges in the region between the substitutions in exon 4 and those in exon 5 (such as the one illustrated) are proposed to produce two fusion genes, one that encodes a pigment that is slightly 'red-shifted' from the normal MWS pigment (protanomalous pigment) and a second that is slightly 'green-shifted' from the normal LWS pigment (deuteranomalous pigment).

responsible for the larger spectral shift that differentiates MWS from LWS pigments. It seems that, except for a crossover between the two exon 4 substitutions (Fig. 4), only a single spectrally intermediate protanomalous hybrid pigment and a single spectrally intermediate deutanomalous pigment could be produced by exchanges upstream of the D_g - D_r marker.

Southern-blot analysis of pigment genes from colour-anomalous subjects does not seem to distinguish between individuals with mild and more extreme forms of anomalous trichromacy. Nathans *et al.*² suggest that different locations of crossing over can produce a variety of anomalous pigments with differing spectral peaks, and that this spectral variation accounts for the differences in colour-discrimination ability among anomalous trichromats. The results presented here indicate otherwise—that both mild and extreme forms of protanomalous colour vision are probably based on the same anomalous photopigment and that, similarly, a second anomalous pigment underlies both mild and extreme forms of deutanomalous colour vision. This conclusion (1) is consistent with recent results from colour matching³, which show that under conditions designed to maximize colour discrimination, both mild and extreme anomalous subjects make the same colour match and thus seem to have the same anomalous pigment; and (2) reinforces the

conclusion reached from psychophysical studies^{3,4,11} that the differences in colour-discrimination ability within each of the two classes of anomalous trichromats cannot be explained by differences in the absorption spectra of the underlying photopigments. □

Received 31 July; accepted 20 October 1989.

- Nathans, J., Thomas, D. & Hogness, D. S. *Science* **232**, 193–202 (1986).
- Nathans, J., Piantanida, T. P., Eddy, R. L., Shows, T. B. & Hogness, D. S. *Science* **232**, 203–210 (1986).
- Nagy, A. L. *J. opt. Soc. Amer.* **72**, 571–577 (1982).
- Pokorny, J. & Smith, V. C. *Color Res. Appl.* **7**, 159–164 (1982).
- Neitz, J. & Jacobs, G. H. *J. opt. Soc. Amer.* **A1**, 1175–1180 (1984).
- Schnapf, J. L., Kraft, T. W. & Baylor, D. A. *Nature* **325**, 439–441 (1987).
- Lehninger, A. L. *Biochemistry* (Worth, New York, 1975).
- Hackett, N. R., Stern, L. J., Chao, B. H., Kronis, K. A. & Khorana, H. G. *J. biol. Chem.* **262**, 9277–9284 (1987).
- Kosower, E. M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1076–1080 (1988).
- Drummond-Borg, M., Deeb, S. & Motulsky, A. G. *Am. J. hum. Genet.* **43**, 675–683 (1988).
- Hurvich, L. M. & Jameson, D. *Mod. Probl. Ophthalmol.* **13**, 200–209 (1974).
- Wyszecki, G. & Stiles, W. S. *Color Science* (Wiley, New York, 1982).
- Dawis, S. M. *Vision Res.* **21**, 1427–1430 (1981).
- DiLella, A. G. & Woo, S. L. C. *Meth. Enzym.* **152**, 199–212 (1987).
- Wahl, G. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 2160–2164 (1987).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).

ACKNOWLEDGEMENTS. We thank D. F. Peters for being a subject in these experiments, K. E. Neitz for her cooperation, and S. K. Fisher for his support. Probes for the X-linked photopigment genes were provided by J. Nathans. This work was supported by the NIH.

Characterization of a naturally processed MHC class II-restricted T-cell determinant of hen egg lysozyme

S. Demotz, H. M. Grey, E. Appella* & A. Sette

Cytel, 11099 North Torrey Pines Road, La Jolla, California 92037, USA

*National Cancer Institute, National Institutes of Health, Building 37, Room 1B04, Bethesda, Maryland 20892, USA

COMPELLING evidence indicates that T cells recognize complexes formed by major histocompatibility complex-encoded molecules and antigenic peptide fragments^{1–3}. This is based largely on the ability of small synthetic peptides to substitute for naturally processed antigen in stimulating T cells. Naturally processed fragments of exogenous antigen are thought to arise by limited proteolytic degradation of native antigen inside acidic compartments of antigen-presenting cells^{3–6}, but until now no physiologically processed antigen has been directly analysed. Here we report the characterization of physiologically processed antigen eluted from mouse class II major histocompatibility complex I-E^d molecules. The antigenic material corresponds to a previously described antigenic determinant of hen egg lysozyme (HEL 107–116) and has a relative molecular mass M_r of about 2,000. HPLC analysis identified at least two or three separate molecular species, suggesting limited, albeit significant, heterogeneity of naturally processed peptides. Finally, under our experimental conditions, it was calculated that a substantial proportion (10–40%) of I-E^d molecules were occupied by these HEL-derived antigenic determinants.

Taking advantage of the high stability of the complexes formed between major histocompatibility complex (MHC) molecules and antigenic peptides, we decided to attempt the isolation and characterization of naturally processed HEL peptides from HEL-pulsed A20-1.11 B lymphoma cells (A20 cells). As a first step, we determined whether complexes containing antigenic determinants could be identified in a preparation of affinity-purified class II MHC molecules. For this purpose, I-E^d molecules were purified from a lysate of HEL-pulsed A20 cells (HEL/I-E^d), and inserted into planar membranes^{8,9}. These structures strongly stimulated interleukin-2 (IL-2) production by the I-E^d-restricted T cell hybridoma 1 H-11.3, for which the

TABLE 1 Isolation of naturally processed antigenic material from antigen-pulsed B lymphoma cells

Source of I-E ^d	Antigen added to purified I-E	IL-2 (U ml) ⁻¹ produced by 1 H-11.3 T cells
HEL-pulsed A20	Nil	1,280
Unpulsed A20	Synthetic HEL 107–116 (20 ng)	20
Unpulsed A20	Synthetic HEL 107–116 (200 ng)	640
Unpulsed A20	Acid-eluted peptides from HEL/I-E ^d	80

A20 cells were grown in RPMI medium supplemented with 5% fetal calf serum, 100 μ g ml⁻¹ streptomycin, 100 U ml⁻¹ penicillin G, and 5×10^{-5} β -mercaptoethanol. A20 cells were pulsed for 24 h with 1 mg ml⁻¹ HEL (Grade I, Sigma). I-E^d molecules were purified from both pulsed and unpulsed cells by affinity chromatography on 14.4.4 anti-I-E^d monoclonal antibody column⁸. To prepare the acid-eluted peptides, I-E^d (350 μ g) was purified from 1.5×10^{10} HEL-pulsed A20 cells (HEL/I-E^d), acetone-precipitated, treated for 30 min at 37 °C with 2.5 M acetic acid, and then loaded on a Sephadex G-10 column equilibrated with acetate buffer¹¹. The peptide material eluted in the void volume was collected and lyophilized twice. Complexes between I-E^d and peptides prepared *in vitro* were obtained by dissolving the lyophilized peptides in 10 μ l phosphate-buffered saline/100 mM Tris-HCl, pH 7.2, supplemented with 1 mg ml⁻¹ sodium azide and a cocktail of protease inhibitors (16 mM EDTA, 2.6 mM 1,10-phenanthroline, 0.15 mM pepstatin A and 2 mM PMSF) (ref. 10), and 10 μ l of a 1–2 mg ml⁻¹ I-E^d preparation. After incubating for 1 d at room temperature, the samples were diluted to 250 μ l with PBS containing 1 mg ml⁻¹ sodium azide, 1% *n*-octyl- β -D-glucopyranoside (Sigma), 175 μ g ml⁻¹ L- α -phosphatidylcholine (Sigma), and 25 μ g ml⁻¹ cholesterol (Sigma). I-E^d-peptide complexes were inserted into planar membranes and tested in a T-cell stimulation assay as described^{8,9}.

sequence HEL 107–116 has previously been identified as the minimal antigenic determinant⁷. Thus, this I-E^d preparation carried naturally processed material corresponding to the HEL 107–116 determinant (Table 1, first row).

The next step in this series of experiments would be to dissociate the antigenic determinants from the I-E^d molecules and biochemically characterize them. For this purpose, it was necessary to devise as sensitive an assay as possible. To decrease the amount of antigenic material required to trigger IL-2 release by the 1 H-11.3 T cell hybridoma, we formed I-E^d/peptide complexes *in vitro*, under conditions already established as optimal¹⁰, and used these complexes, inserted into planar mem-

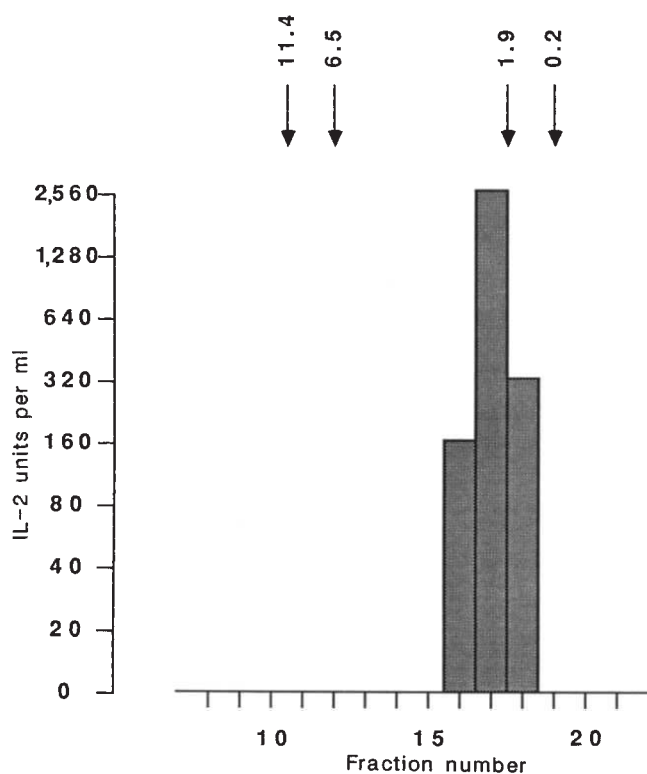


FIG. 1 Gel filtration analysis of a naturally processed I-E^d-restricted determinant of lysozyme.

METHODS. Acid-eluted material obtained from 600 μ g HEL/I-E^d complexes (Table 1) was loaded on a Sephadex G-50 column (1.5 $\times$ 90 cm) equilibrated with 60 mM ammonium acetate buffer, pH 4.8 (ref. 11). Fractions (8 ml) were collected and lyophilized twice. One-fifth of each fraction was tested as described in the legend to Table 1. The positions of molecular weight markers are indicated (*M_r* 11.4 K, cytochrome *c*; 6.5 K, aprotinin; 1.9 K, Y-HEL 105-120; 0.2 K, Cr₂O₇²⁻). In this experiment, I-E^d complexes prepared with 20 ng and 200 ng synthetic HEL 107-116 peptide stimulated 1 H-11.3 T cells to produce 1,280 U ml⁻¹ IL-2 and >2,560 U ml⁻¹ IL-2, respectively.

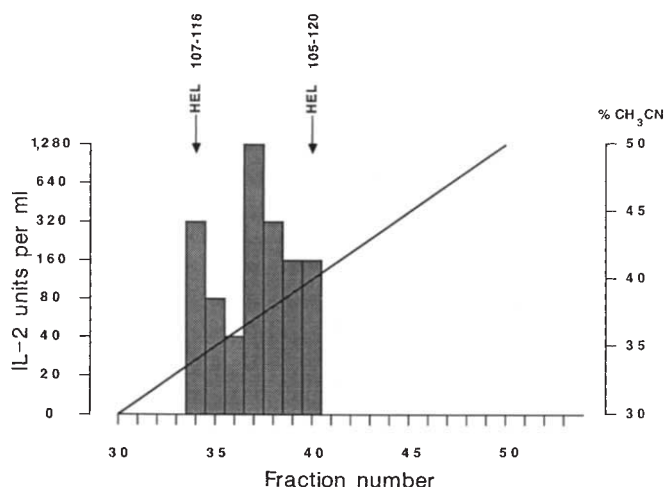


FIG. 2 Reverse-phase HPLC analysis of a naturally processed I-E^d-restricted determinant of lysozyme.

METHODS. Acid-eluted material obtained from 350 μ g HEL/I-E^d complexes (Table 1) was separated on a μ -Bondapak C-18 reverse-phase HPLC column (Waters System) by forming a linear gradient of 0.08% trifluoroacetic acid in water and 0.08% trifluoroacetic acid in acetonitrile (1% per min). Fractions (1 ml) were collected, lyophilized, and tested as described in the legend to Table 1. The elution positions of the synthetic HEL 107-116 and HEL 105-120 peptides are indicated.

branes, in the T cell hybridoma stimulation assay^{8,9}. Following this approach, we could detect as little as 20 ng HEL 107-116 peptide (Table 1), whereas, if directly added to antigen-presenting cells, 250 ng were required to reach similar levels of T-cell stimulation (ref. 7, and data not shown).

A crude preparation of peptides bound to HEL/I-E^d complexes was obtained by acid treatment of acetonitrile-precipitated HEL/I-E^d material, followed by Sephadex G-10 chromatography¹¹. As assessed by SDS-gel electrophoresis, this step completely depleted the acid-eluted HEL/I-E^d peptide preparation of I-E^d molecules (data not shown). The acid-eluted HEL/I-E^d peptides were then incubated with I-E^d molecules purified from A20 cells grown in the absence of any HEL material, and the putative complexes were inserted into planar membranes as already described. These antigen-presenting structures stimulated IL-2 release from 1 H-11.3 T cells, indicating that peptide material containing the HEL 107-116 sequence had been eluted from the I-E^d isolated from cells incubated with HEL, and that these peptides could be rebound to I-E^d molecules (Table 1, last row). Similar data were obtained in the case of an I-A^d-restricted and ovalbumin 323-339 peptide-specific T cell hybridoma (data not shown).

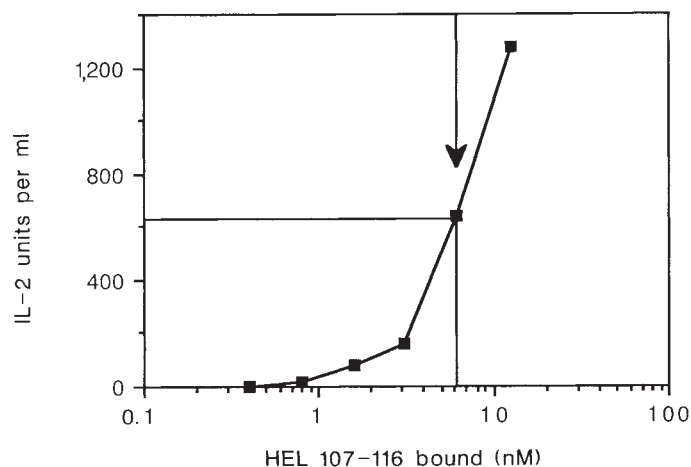
To initiate the characterization of this acid-eluted HEL/I-E^d material, the molecular weight was estimated by gel filtration on a Sephadex G-50 column. Each fraction was tested for its capacity to produce, when recombined to I-E^d, complexes antigenically active for 1 H-11.3 T cells. Stimulatory material was eluted in a single homogeneous peak, with an apparent relative molecular mass of $\sim$ 2,000 (*M_r* 2K) (Fig. 1).

To further characterize the naturally processed HEL-derived material, an acid-eluted peptide preparation was subjected to reverse-phase HPLC and each fraction was tested for its capacity, when incubated with I-E^d, to form T-cell stimulatory complexes. Material antigenic for 1 H-11.3 T cells was eluted between 34-40% acetonitrile. The activity profile suggested that at least two or three peptide species were produced by A20 cells during processing of HEL molecules (Fig. 2). The synthetic peptides HEL 107-116 and HEL 105-120 eluted at 34% and 40% acetonitrile, respectively, suggesting that the HEL fragments generated by A20 cells are not grossly dissimilar from these peptides.

To evaluate the possibility that limited degradation of the antigenic material occurred during the isolation procedure, a radioiodinated synthetic peptide corresponding to [¹²⁵I]Y-HEL 105-120 was purified as described. No degradation could be detected by reverse-phase HPLC, indicating that the antigenic HEL peptides that we characterized were representative of those peptides produced by processing in A20 cells (data not shown).

This assay allowed us to estimate directly the proportion of I-E^d molecules on HEL-pulsed A20 cells that were occupied by peptides containing the HEL 107-116 determinant detected by 1 H-11.3. We therefore compared the antigenicity of an I-E^d preparation from HEL-pulsed A20 (HEL/I-E^d) with synthetic HEL 107-116/I-E^d complexes prepared *in vitro*. The amount of synthetic HEL 107-116 peptide bound to I-E^d molecules was first determined by Scatchard analysis, as previously described¹⁰. Thus it was calculated that under saturating concentrations of HEL 107-116 peptide (>10 μ g ml⁻¹), 7.5% of I-E^d molecules were occupied by this peptide. Then various amounts of synthetic HEL 107-116/I-E^d complexes were inserted into planar membranes, and the capacity to stimulate 1 H-11.3 T cells was measured for each (Fig. 3). The total amount of I-E^d molecules inserted into each planar membrane sample was kept constant by adding I-E^d molecules purified from unpulsed A20 cells, to obtain a direct correlation between the T-cell response and the density of antigenic complexes borne by the planar membranes. The response to the synthetic HEL 107-116 peptide/I-E^d complexes was then compared with naturally processed HEL. It was found that HEL/I-E^d complexes inserted at 16 nM into planar membranes stimulated 1 H-11.3 T cells to produce 640 U ml⁻¹

FIG. 3 Occupancy of I-E^d molecules purified from HEL-pulsed B lymphoma cells by HEL 107-116-containing determinants. To generate a standard curve, various concentrations of synthetic HEL 107-116 peptide bound to I-E^d molecules and inserted into planar membranes were tested for their ability to trigger 1 H-11.3 T cell stimulation. Then a known concentration of I-E^d purified from HEL-pulsed A20 cells (16 nM) was inserted into planar membranes and used to stimulate 1 H-11.3 T cells. An equivalent stimulation to 6 nM of synthetic peptide HEL 107-116/I-E^d complexes was obtained. METHODS. The concentration (nM) of purified I-E^d molecules able to bind antigen was determined by Scatchard analysis¹⁰. The amount of synthetic HEL 107-116 peptide necessary to fully saturate the same I-E^d molecules was measured by testing a dose range of synthetic HEL 107-116 peptide for its capacity to abrogate binding of ¹²⁵I-labelled λ -repressor 12-26 peptide¹⁰. I-E^d molecules (0.5 mg ml⁻¹) were then incubated with 1 mg ml⁻¹ synthetic HEL 107-116 peptide (30-fold excess over the saturating dose), using the same conditions as in the binding assay (see Table 1 legend). After the incubation, I-E^d molecules were recovered from free peptide by passage over a Sephadex G-50 column and inserted into planar membranes^{8,9}. The total amount of I-E^d molecules inserted was kept constant (10–25 μ g ml⁻¹) by adding I-E^d molecules purified from unpulsed A20 cells. In parallel, I-E^d molecules were prepared from HEL-pulsed A20 cells, inserted into planar membranes, and tested for 1 H-11.3 T cell stimulation. As before, I-E^d concentration was adjusted at $\sim 25 \mu$ g ml⁻¹ by the addition of I-E^d purified from unpulsed A20 cells.



IL-2. A similar T-cell activation was reached with 6 nmol synthetic HEL 107-116 peptide bound to I-E^d molecules (Fig. 3). Assuming that HEL 107-116/I-E^d complexes were as stimulatory as naturally processed HEL/I-E^d complexes, these data would indicate that 38% of the I-E^d molecules expressed by HEL-pulsed A20 cells were occupied by HEL 107-116-like determinants.

An independent estimate of the extent of occupancy was based on the amount of antigenic peptides eluted from an HEL/I-E^d preparation after its separation by G-50 gel filtration separation. In this experiment, antigenic material obtained from 2 nmoles of I-E^d yielded T-cell stimulatory activity corresponding to 2,000–3,000 U ml⁻¹ IL-2 (Fig. 1). Similar levels of IL-2 production were attained using ~ 0.2 nmol synthetic HEL 107-116 peptide. So by this calculation it seems that $\sim 10\%$ of the I-E^d molecules purified from HEL-pulsed A20 cells carried HEL 107-116 determinants. This value is probably a minimal estimate as the G-50 fractions would also be expected to contain non-HEL-derived I-E^d-binding peptides which would tend to compete with and lower the efficiency of binding of the HEL peptides to I-E^d (ref. 11).

In conclusion, the preparation of complexes between MHC molecules and antigenic peptides allowed us to characterize a naturally processed determinant of HEL. The heterogeneity of the antigenic peptides produced by A20 cells may be due to cleavages at different positions in the HEL sequence, reflecting the use of either different proteases or proteases displaying a broad substrate specificity.

Quantitation of the proportion of I-E^d from HEL-pulsed A20 cells bearing HEL 107-116-containing determinants gave surprisingly high estimates (10–40%). These data can be rationalized, at least in part, by considering that HEL represented, under our experimental conditions, about one-third of the total extracellular protein (1 mg ml⁻¹ HEL and 2–3 mg ml⁻¹ of fetal calf serum proteins). Furthermore, the HEL 107-116 peptide region is the chief I-E^d binding site of HEL (ref. 7). We cannot exclude the possibility, however, that there are covalent modifications of antigenic material which increase its antigenic potency¹². Direct sequencing of naturally processed antigenic peptides should solve these issues. □

- Chesnut, R. W., Colon, S. M. & Grey, H. M. *J. Immun.* **129**, 2382–2387 (1982).
- Ziegler, K. & Unanue, E. R. *J. Immun.* **127**, 1869–1875 (1981).
- Streicher, H. Z., Berkower, I. J., Busch, M., Gurd, F. R. N. & Berzofsky, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6831–6835 (1984).
- Adorini, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 5181–5185 (1988).
- Coeshott, C. M. *et al. J. Immun.* **136**, 2832–2838 (1986).
- Buus, S., Sette, A., Colon, S. M., Jenis, D. M. & Grey, H. M. *Cell* **47**, 1071–1077 (1986).
- Sette, A., Buus, S., Colon, S., Miles, C. & Grey, H. M. *J. Immun.* **142**, 35–40 (1989).
- Buus, S., Sette, A., Colon, S. M. & Grey, H. M. *Science* **242**, 1045–1047 (1988).
- Falo, L. D. Jr., Benacerraf, B. & Rock, K. L. *Proc. natn. Acad. Sci. U.S.A.* **83**, 6994–6997 (1986).

ACKNOWLEDGEMENTS. We thank Joyce Joseph for secretarial assistance. Stéphane Demotz was supported by the Swiss National Fund for Scientific Research.

Sequence and expression of a frog brain complementary DNA encoding a kainate-binding protein

Keiji Wada*, Claude J. Dechesne*, Shunichi Shimasaki†‡, Ron G. King§, Kiyoshi Kusano||, Andres Buonanno||, David R. Hampson*‡, Carl Banner§, Robert J. Wenthold*# & Yoshihiro Nakatani§

* Laboratory of Molecular Otology, NIDCD, § Laboratory of Molecular Biology, || Laboratory of Neurochemistry, NINDS and ¶ Laboratory of Developmental Neurobiology, NICHD, National Institutes of Health, Bethesda, Maryland 20892, USA

† Neuroendocrinology Laboratory, The Salk Institute for Biological Studies, La Jolla, California 92037, USA

EXCITATORY amino acids (EAAs) are important neurotransmitters in the vertebrate central nervous system¹. Electrophysiological and ligand-binding studies indicate that at least three different receptor subtypes for EAAs exist—N-methyl-D-aspartate, kainate and quisqualate receptor subtypes—on the basis of the preferred agonist of the receptors^{1,2}. We recently purified a kainate-binding protein (KBP) from frog (*Rana pipiens berlandieri*) brain by domoic acid (a high-affinity kainate analogue) affinity

Received 19 July; revised 18 September; accepted 23 October, 1989.

- Shimonkevitz, R., Colon, S., Kappler, J. W., Marrack, P. & Grey, H. M. *J. Immun.* **133**, 2067–2074 (1984).
- Babbitt, B., Allen, P., Matsueda, G., Haber, E. & Unanue, E. *Nature* **317**, 359–361 (1985).
- Buus, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 3968–3971 (1986).

To whom correspondence should be addressed at: Neurochemistry Section, Laboratory of Molecular Biology, Bldg 36, Rm 5D08, National Institutes of Health, Bethesda, MD 20892, USA.

‡ Present address: Department of Molecular Endocrinology, The Whittier Institute, La Jolla, CA 92037, USA (S.S.); Faculty of Pharmacy, University of Toronto, Toronto, Ontario M5S 2S2, Canada (D.R.H.).

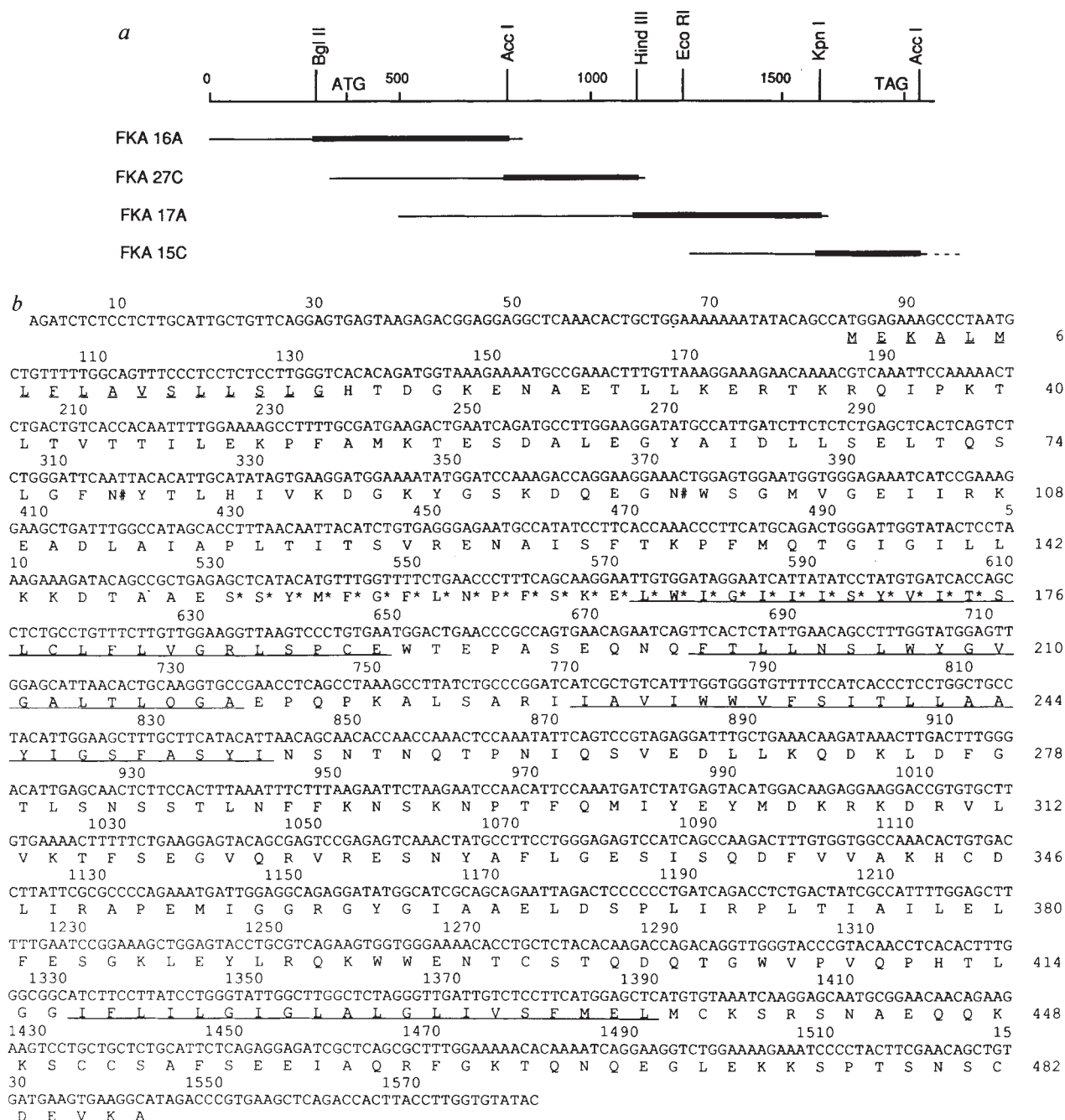


FIG. 1 *a*, Restriction enzyme map of cDNA clones encoding the KBP. Thick lines represent regions that are used to construct an expressible cDNA containing an entire open-reading frame. Broken line shows the regions not sequenced. Positions of the initiator methionine codon and stop codon of an open reading frame are shown by ATG and TAG, respectively. *b*, Nucleotide and deduced amino-acid sequence (single-letter code) of the KBP cDNA. Only the sequences used for the expression studies are shown. Nucleotides and amino acids are numbered in order, starting with the first residues. Amino-acid residues with asterisks show the chemically obtained KA2 sequence. Individually underlined amino-acid residues show the predicted signal peptide sequence¹⁰. Underlined sequences indicate the putative transmembrane domains predicted by regional hydropathy analysis^{16,17}. Potential *N*-linked glycosylation sites¹¹ are indicated by #.

METHODS. KBP was isolated from frog brain by domoic-acid affinity chromatography³. The protein was further purified by SDS-PAGE. The main band migrating to a position corresponding to an *M_r* of 48,000 was excised, digested with *S. aureus* V8 protease in a stacking gel, and fractionated by electrophoresis on a 15% SDS-polyacrylamide gel. The gel was electroblotted onto an Immobilon membrane (Millipore, Bedford, Massachusetts). After staining with Coomassie blue, pieces containing the main digested fragments were excised and microsequenced. One long amino-acid sequence (KA2) was obtained from one of the fragments (single-letter code):

SSYMFGLNPFSLXIGIISYVIT, where X is an unidentified residue. Sense and antisense 20-base oligonucleotides were made using all possible codon combinations encoding the sequence near the N and C termini of KA2, respectively: 5'-TAYATGTTGGNTTYTNA-3', 5'-ACRTANSWDATDAT-DATNCC-3'. R, A or G; S, C or G; W, A or T; Y, C or T; D, A or G or T; and N, A or C or G or T. PCR (ref. 8) was performed using the oligonucleotides as primers and frog liver genomic DNA as a template. Frog liver genomic DNA was prepared by a standard method²⁴. Annealing reactions were performed at 50 °C for 2.5 min, followed by a 3.5-min extension at 72 °C and 1-min denaturation at 94 °C. After 35 rounds of amplification, PCR-extended fragments of 65-bp were purified, subcloned into the *Sma* site of the plasmid pGEM 3Z (Promega, Madison, Wisconsin) and sequenced²⁵. Based on this resultant sequence, 65-base oligonucleotides were prepared and used to screen a frog brain cDNA library (Clontech, Palo Alto, California) in λ gt11 vector: 5'-TACATGTTCCGGTTTCTGAACCCCTTCTCCAAGGAGCTGTGATCGG-CATCATCATAACCTACGT-3'. Hybridization and washing of filters were carried out according to a standard method²⁴. The isolated cDNAs were subcloned into the plasmid pBluescript KS (+) (Stratagene, La Jolla, California) and sequenced²⁵. Some nucleotide sequence differences were found between the PCR product and the KBP cDNA. These could have been due to replication errors of the reverse transcriptases and DNA polymerases used, and/or to genetic polymorphism among frogs.

chromatography³, and showed that the kainate-binding activity was associated with a protein of relative molecular mass 48,000 (M_r 48 K)^{3,4}. The pharmacological properties and the anatomical distribution of KBP were consistent with those of a kainate receptor-ionophore complex³⁻⁵. We have now isolated a complementary DNA encoding KBP of M_r 48 K. The deduced amino-acid sequence of the KBP has similar hydrophobic profiles to those found in other ligand-gated ion channel subunits⁶, and shows some amino-acid sequence similarities to the corresponding regions of brain nicotinic acetylcholine receptor subunits⁷. Localization of the KBP messenger RNAs by *in situ* hybridization histochemistry is compatible with the results of immunohistochemistry and receptor autoradiography studies⁵. COS-7 cells transfected with the cDNA encoding the KBP show high-affinity kainate-binding activity with pharmacological properties similar to those of the biochemically purified KBP (ref. 3). These results provide the first molecular characterization of an EAA-binding site and raise the possibility that the KBP cDNA encodes a ligand-binding subunit of a kainate receptor-ionophore complex.

A sequence of 26 amino-acid residues (referred to as KA2), as well as other shorter sequences, was obtained from the peptide fragments of the purified KBP which had been digested with *Staphylococcus aureus* V8 protease. Because none of the peptide sequences contained a suitable region to make oligonucleotide probes with little degenerativity, we used the polymerase chain reaction (PCR) procedure⁸ to obtain a precise nucleotide sequence for KA2 (see Fig. 1 legend for experimental details). A fragment of genomic DNA of the expected length was

amplified from a pool of frog liver genomic DNA. Sequence analysis of the fragment revealed a nucleotide sequence that encoded the KA2 sequence between the two regions used to prepare the primers. The revealed nucleotide sequences and the primer sequences in the PCR product were used to make a long oligonucleotide probe to screen a frog brain cDNA library in λ gt11 vector. From 7×10^5 recombinants, four clones were isolated and used to obtain longer cDNA clones. Four overlapping longer clones (FKA16A, FKA27C, FKA17A and FKA15C) were obtained from a total of 1.4×10^6 recombinations and studied further (Fig. 1a).

Sequence analysis of these clones showed a 1,461-base-pair (bp) open-reading frame potentially encoding 487 amino-acid residues (Fig. 1b). The sequence around the predicted initiator methionine codon (ATG) agrees with the consensus sequence described by Kozak⁹. The deduced amino-acid sequence shows a hydrophobic segment at the N terminus, which probably encodes a signal peptide. The N terminus of the mature protein was predicted by the method of von Heijne¹⁰. The proposed mature protein is preceded by a signal peptide of 17 residues and composed of 470 amino-acid residues with a calculated M_r of 52.5 K. Two potential N-linked glycosylation sites¹¹ are found at position 78 and 96 in the putative N-terminal extracellular segment. The amino-acid sequence of the predicted mature protein contains the KA2 sequence at residues 150–175. Other shorter amino-acid sequences obtained by the microsequencing of the V8 protease-treated fragments are also found at residues 57–66, 62–74, 110–125 and 353–366.

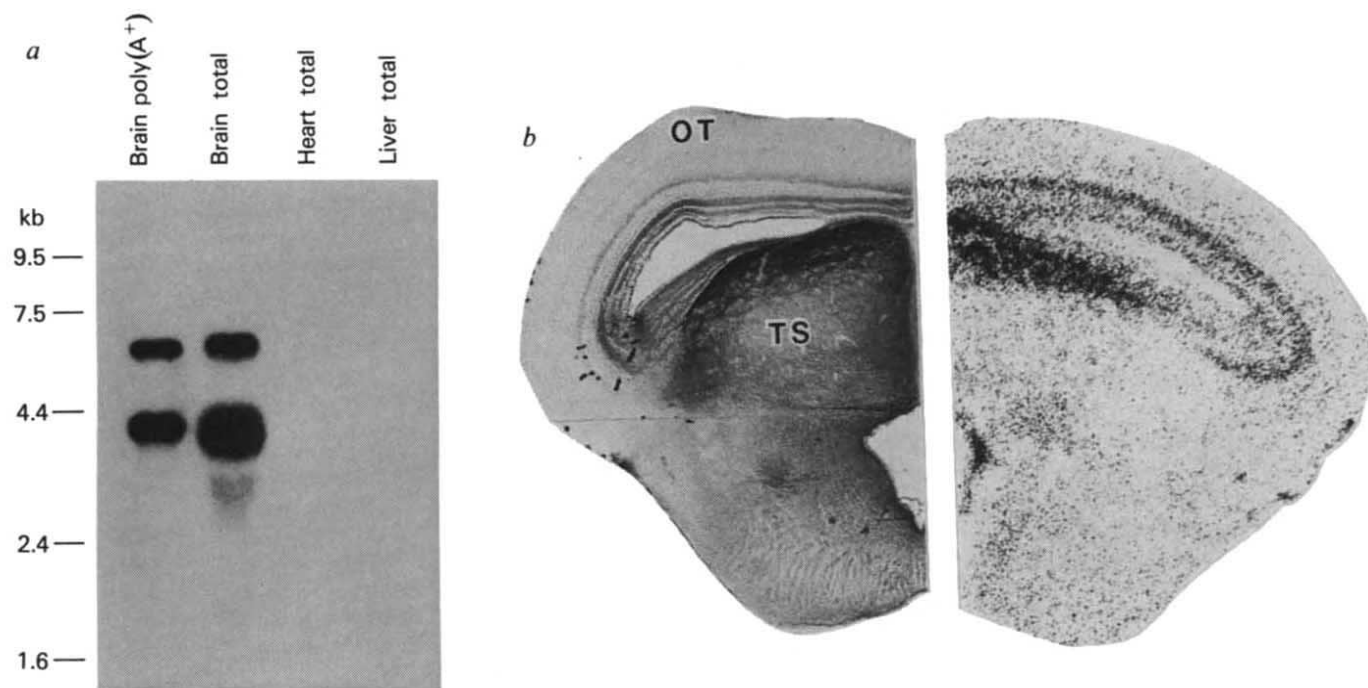


FIG. 2 *a*, Northern-blot analysis of the KBP mRNA. *b*, Comparison of labelling pattern in frog brain with immunostaining of KBP (left) and *in situ* hybridization histochemistry of KBP mRNA (right). Coronal sections of mesencephalon and optic tectum are presented. TS, Torus semicircularis; OT, Optic tectum.

METHODS. *a*, Frog brain total and poly(A)⁺ RNA, and frog heart and liver total RNAs were prepared using a standard method²⁶. Each total RNA (10 μ g) and 0.5 μ g poly(A)⁺ RNA were electrophoresed on a 1% formaldehyde-agarose gel and blotted onto a nitrocellulose membrane. Hybridization was carried out in a buffer containing $5 \times$ SSC and 50% formamide at 42 °C. The pBluescript plasmid containing the expressible KBP cDNA sequence was used as a probe. The membrane was washed with $2 \times$ SSC at 55 °C and exposed overnight. The same pattern of signals was obtained after a very high stringency wash with $0.1 \times$ SSC at 70 °C. *b*, Immunohistochemistry was performed as previously described⁵. After incubating with the monoclonal

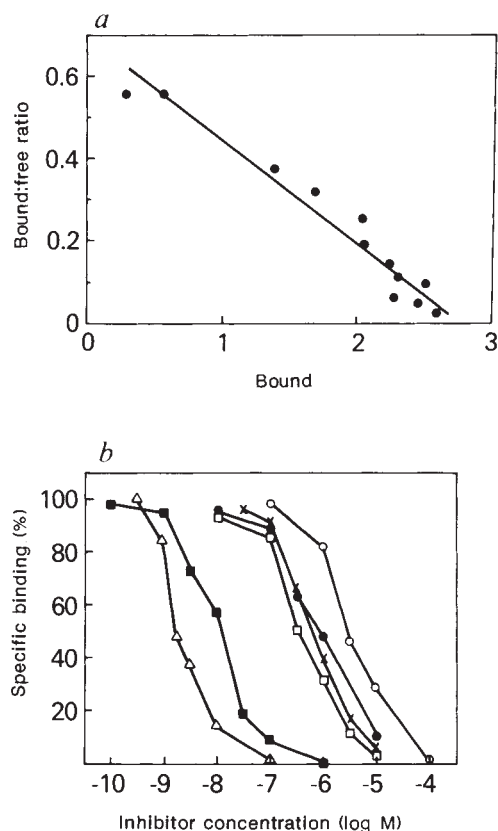
antibody, KAR-A1 (refs 4, 5), immunostaining was visualized with avidin-biotin-peroxidase complex. *In situ* hybridization histochemistry was performed on freshly frozen frog brain sections. Sections (15 μ m) were cut on a cryostat and mounted onto gelatin- and poly-L-lysine-coated slides. After fixation in phosphate buffer containing 4% paraformaldehyde for 10 min at 4 °C followed by rinsing in $2 \times$ SSC, sections were hybridized with ³⁵S-labelled oligonucleotide probes overnight in buffer containing $5 \times$ SSC and 50% formamide at 37 °C. Two 48-base oligonucleotides complementary to the sequences at residues 421–468 and 469–516 (Fig. 1) were used as probes. After hybridization, slides were washed in $0.5 \times$ SSC at 40 °C, dehydrated and exposed to X-ray films for 20 days. For higher resolution studies, slides were then dipped in Kodak NTB-3 nuclear emulsion and exposed for 3 weeks. An oligonucleotide probe encoding the sense strand (residues 421–468) was used as a negative control. Both antisense oligonucleotide probes gave the same results.

FIG. 3 a, Scatchard analysis of ^3H -labelled kainate binding to transfected COS-7 cell membranes. Binding was performed at concentrations of 0.5–100 nM. Data were analysed using the LIGAND program¹³, and found to fit a single-site model. A K_d of 5.6 ± 0.7 nM (mean $\pm$ s.d., $n=3$) was obtained. Values for the maximum number of binding sites were 2.8 and 3.1 pmol per mg protein for two separate batches of cells. Bound are expressed as pmol ^3H -labelled kainate per mg protein and bound:free ratio is expressed as pmol ^3H -labelled kainate per mg protein per nM. b, Inhibition profiles of various ligands for ^3H -labelled kainate binding to the transfected cells. Values of 50% inhibitory concentration IC_{50} (mean $\pm$ s.d., $n=3$) were 2.7 ± 0.7 nM for domoic acid (Δ), 13.2 ± 0.5 nM for kainate ($\blacksquare$), 0.59 ± 0.35 μM for quisqualate ($\square$), 0.80 ± 0.35 μM for glutamate ($\bullet$), 0.83 ± 0.13 μM for 6-cyano-7-nitroquinoxaline-2,3-dione ($\times$), and 4.1 ± 1.2 μM dihydro-kainate ($\circ$).

METHODS. An expressible cDNA containing an entire open-reading frame was constructed as follows (see also Fig. 1). The cDNA insert of FKA 27C was subcloned into *EcoRI* site of the plasmid bBluescript KS(+) (Stratagene). The *HindIII*-*KpnI* fragment of FKA 17A, the *KpnI*-blunt-ended *AccI* fragment of FKA 15C and the *HindIII*-*HincII* fragment of the plasmid containing FKA 27C cDNA were ligated to yield FKA 271715. The *BglII*-*AccI* fragment of FKA 16A and the *BamHI*-*AccI* fragment of the FKA 271715 were ligated to yield FKA 16271715. The *XbaI*-*Apal* fragment of the plasmid FKA 16271715, containing the expressible cDNA, was subcloned under the SV40 early promoter of a modified form of the plasmid, pcD2 (ref. 12). The modified pcD2 plasmid was made by the introduction of multiple cloning sites at *PstI* and *BamHI* sites of the original plasmid pcD2 vector. The sequence encoding the introduced multiple cloning sites was 5'-TGCAGTCTAGACCGGGAATTCGGGCCCGGATC-3'. COS-7 cells (5×10^5 per 10-cm plate) were cultured overnight and transfected with 20 μg plasmid pcD2 DNA containing the expressible FKA cDNA insert by a calcium phosphate method¹². After 24 h, the medium was changed and plates were incubated for an additional 4 days. Cells were collected in distilled water, centrifuged at 54,000g for 20 min, resuspended with a polytron in distilled water, and centrifuged at 54,000g for 20 min. Membrane fractions were frozen as pellets and stored at -70°C . For binding studies, frozen membranes were resuspended with a polytron in 50 mM Tris-HCl, pH 7.0, and centrifuged at 54,000g for 20 min. The membranes were washed two additional times and subjected to ^3H -labelled kainate binding as previously described³. Untransfected COS-7 cells had no binding activity for ^3H -labelled kainate. For displacement studies, inhibitors were incubated with membranes for 10 min before addition of ^3H -labelled kainate at a final concentration of 5 nM.

Northern blot analysis with KBP cDNA probe showed that two principal transcripts of 6 and 4 kilobases were found in frog brain (Fig. 2a). *In situ* hybridization histochemistry on frog brain sections demonstrated that the pattern of expression of the KBP transcripts was similar to those of the biochemically purified KBP and ^3H -labelled kainate-binding sites (Fig. 2b, and ref. 5). The KBP transcripts were abundantly expressed in the torus semicircularis and the optic tectum. Higher resolution study revealed that the transcripts were distributed in layers 2, 4 and 6 of the optic tectum. Dendrites from neurons in these layers project into layers 3 and 5, where immunostaining was apparent⁵.

An expressible cDNA containing the entire open-reading frame was constructed using appropriate restriction-enzyme-digested fragments from the four positive clones in Fig. 1a, and subcloned under the SV40 early promoter in a modified form of the plasmid pcD2 vector¹². The plasmids were transfected into COS-7 cells by a calcium phosphate method¹². Membrane fractions of the transfected cells were prepared 4 days after transfection and assayed for ^3H -labelled kainate-binding activity. Scatchard analysis¹³ indicated that the membrane fraction of the transfectants contained a single class of ^3H -labelled kainate-binding sites with a dissociation constant (K_d) of 5.6 nM (Fig. 3a). Purified frog brain KBP has both a high ($K_d = 5.5$ nM)- and a low ($K_d = 34$ nM)-affinity binding site for ^3H -labelled kainate³. Lack of the low-affinity binding site in the transfectants could indicate the existence of another subunit in the purified frog brain protein. Alternatively, different post-translational modifications between frog brain and transfected COS-7 cells could explain the absence. Displacement of ^3H -labelled kainate binding with selective ligands showed similar patterns to those found in the biochemically purified KBP (Fig. 3b, and ref. 3).



Immunoblotting analysis of the membrane preparation using polyclonal antisera against the purified frog KBP revealed that the isolated KBP cDNA produced a protein with a migration rate on SDS-PAGE similar to that of the purified KBP of M_r 48 K (data not shown).

Although the expressed protein in transfected cells could bind kainate, injection of mRNA, transcribed *in vitro* using the bacterial RNA polymerase from the cDNA encoding the entire KBP, did not produce a detectable electrophysiological response to perfused kainate in *Xenopus laevis* oocytes by a two-micro-electrode voltage-clamp method. In preliminary experiments, co-injection of oligonucleotides complementary to the KBP cDNA sequence with frog brain poly(A)⁺ RNA^{14,15} did not significantly repress an electrophysiological response to kainate compared with control.

Analysis for regional hydrophobicity^{16,17} of the deduced amino-acid sequence for the KBP revealed four putative transmembrane domains. The distribution and spacing of the domains are similar to those seen in various subunits of ligand-gated ion channels, including nicotinic acetylcholine receptors (nAChRs)⁷, glycine receptor¹⁴ and γ -aminobutyric-acid receptors¹⁸ (Fig. 4a). This hydrophobic profile is one of the common structural features presumed to be important for the function of the ligand-gated ion channels⁶. But neither the two cysteine residues in the predicted N-terminal extracellular segment nor the proline residue in the first proposed transmembrane domain, which are other important structural features conserved in ligand-gated ion channel subunits⁶, is found in the KBP. The KBP has a shorter N-terminal and a longer C-terminal hydrophilic segment than do the ligand-gated ion channel subunits.

Although homology search in a sequence bank (NBRF) revealed no strong similarities to known proteins, some amino-

acid sequence similarity to neuronal nAChR subunits⁷ could be found in particular segments, especially in the proposed fourth transmembrane domain (Fig. 4b). The sequence alignment of the KBP and neuronal nAChR α_2 -subunit shows that the serine residue in the proposed second transmembrane domain is conserved (Fig. 4b). The equivalent residue in *Torpedo* and muscle nAChR subunits is important for ion transport^{19,20}. The KBP has negatively charged amino-acid residues around the second putative transmembrane domain. Negatively charged amino-acid residues seen in the homologous region of the *Torpedo* nAChR subunits have been proposed to form anionic rings, which are important determinants of the cation transport rates²¹. Like the nAChRs, kainate receptors are known to be cation-transfer ion channels¹.

The results presented here show that the KBP cDNA encodes a kainate-binding protein that cannot form a functional channel when expressed alone. There are two general interpretations for this finding. The first is that the KBP cDNA does not encode a neurotransmitter receptor subunit, but rather an uncharacterized kainate-binding molecule, which may or may not be related to EAA neurotransmission. The preliminary results of the complementary oligonucleotide blocking experiments and the lack of certain common structural features in the KBP, which are

presumed to be important for channel function and found in all ligand-gated ion channel subunits sequenced so far, could support this interpretation. Because KBP is the principal kainate-binding protein in frog brain³⁻⁵, this interpretation would imply that most or all of the kainate-binding sites are not associated with a kainate receptor-ionophore complex. The fact that the KBP has some structural properties of known ligand-gated ion channel subunits indicates that such a binding protein is evolutionarily and functionally related to a receptor-ionophore complex.

The second interpretation is that the KBP cDNA encodes a subunit of a kainate-gated ionophore that requires an additional subunit(s) for physiological activity. The requirement of multiple subunits to produce a functional receptor-ionophore complex is well documented for nAChRs (refs 7, 22, 23). Injection of *Torpedo* nAChR α -subunit mRNA alone produces a protein with ligand-binding activity but no channel activity in *Xenopus* oocytes²². Furthermore, the following observations in the present study support the idea that the KBP is related to a functional receptor-ionophore complex: (1) COS-7 cells transfected with the KBP cDNA could bind kainate; (2) hydrophobic profiles of the KBP were similar to those of other ligand-gated ion channel subunits; (3) some sequence similarities to nAChR

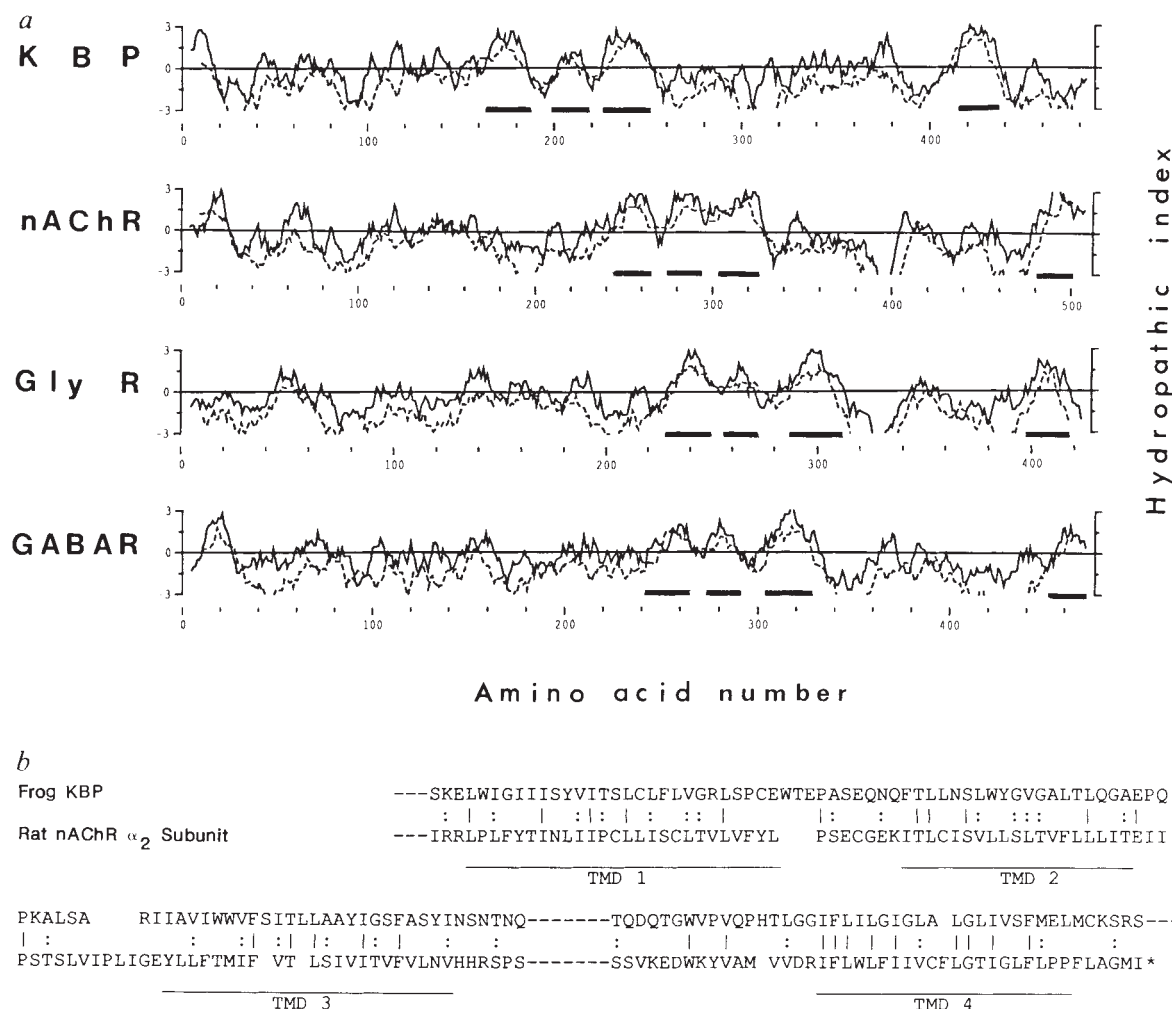


FIG. 4 *a*, Comparison of hydropathy profiles of the KBP, α_2 -subunit of rat nAChR⁷, strychnine-binding subunit of rat glycine receptor (GlyR)¹⁴ and β_1 subunit of bovine γ -aminobutyric acid receptor (GABAR)¹⁸. Scores of hydropathy were computed according to Kyte and Doolittle¹⁶ (solid line) and Engelman, Steitz and Goldman¹⁷ (broken line) with window sizes 9 and 20, respectively. Solid bars indicate proposed transmembrane domains.

b, Alignment of the amino-acid sequence of the KBP and nAChR α_2 subunit⁷ in the putative transmembrane domains. Identical residues are indicated by vertical lines. Colons indicate conservative substitution of amino-acid residues²⁷. Gaps of sequence are inserted to achieve maximum similarities. Dashes show the sequences not compared. TMD, Transmembrane domains.

subunits were found in certain segments; and (4) there could be an additional homologous subunit(s) forming a functional receptor together with the KBP. This possibility is supported by the fact that two bands were detected by northern blotting at high stringency, and that multiple proteins at positions corresponding to a M_r of 48,000 on an SDS-polyacrylamide gel were recognized by immunostaining using polyclonal and monoclonal antibodies against the KBP (refs 3, 4). Furthermore, the KBP cDNA encodes a protein that produces only a high-affinity binding site for kainate in COS-7 cells. Although the lack of the previously observed lower-affinity site³ could be due to different post-translational modifications between frog brain and transfected COS-7 cells, it raises the possibility that a second

protein, which contains the low-affinity binding site, is necessary for a physiologically functional receptor.

This study provides the first molecular characterization of a component of a kainate receptor-binding system that could be involved in EAA neurotransmission. Recent evidence has shown that EAAs are important neurotransmitters in the central nervous system and could be involved in neurotoxicity and several other neurological disorders^{1,2}. This provides an impetus to characterize their receptors and related proteins. Further studies focused on isolating and expressing additional subunits of KBP should help elucidate the precise role of the protein and its relationship to neurotransmission and kainate neurotoxicity. □

Received 12 September; accepted 12 October 1989

- Monaghan, D. T., Bridge, R. J. & Cotman, C. W. *A. Rev. Pharmac. Tox.* **29**, 365–402 (1989).
- Foster, A. C. & Fagg, G. *Brain Res. Rev.* **7**, 103–164 (1984).
- Hampson, D. R. & Wenthold, R. J. *J. biol. Chem.* **263**, 2500–2505 (1988).
- Hampson, D. R., Wheaton, K. D., Dechesne, C. J. & Wenthold, R. J. *J. biol. Chem.* **264**, 13329–13335 (1989).
- Dechesne, C. J. *et al. J. Neurosci.* (in the press).
- Barnard, E. A., Darlison, M. G. & Seeburg, P. *Trends Neurosci.* **10**, 502–509 (1987).
- Wada, K. *et al. Science* **240**, 330–334 (1988).
- Saiki, R. K. *et al. Science* **239**, 487–491 (1988).
- Kozak, M. *Nucleic Acids Res.* **12**, 857–872 (1984).
- von Heijne, G. *Nucleic Acids Res.* **14**, 4683–4690 (1986).
- Marshall, R. D. *Biochem. Soc. Symp.* **40**, 17–26 (1974).
- Chen, C. & Okayama, H. *Molec. cell. Biol.* **7**, 2745–2752 (1987).
- Munson, P. J. & Rodbard, D. *Analyt. Biochem.* **98**, 39–45 (1980).
- Grenningloh, G. *et al. Nature* **328**, 215–220 (1987).
- Sumikawa, K. & Miledi, R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1302–1306 (1988).
- Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
- Engelman, D. M., Steitz, T. A. & Goldman, A. *Rev. Biophys. biophys. Chem.* **15**, 321–353 (1986).

- Schofield, P. R. *et al. Nature* **328**, 221–227 (1987).
- Giraudat, J. *et al. Biochemistry* **26**, 2410–2418 (1987).
- Leonard, R. J., Labarca, C. G., Charriet, P., Davidson, N. & Lester, H. A. *Science* **242**, 1578–1581 (1988).
- Imoto, K. *et al. Nature* **335**, 645–648 (1988).
- Kurosaki, T. *et al. FEBS Lett.* **214**, 253–258 (1987).
- Boulter, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 7763–7767 (1987).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
- Okayama, H. *et al. Meth. Enzym.* **154**, 3–28 (1987).
- Dayhoff, M. O., Schwartz, R. M. & Orcutt, B. C. in *Atlas of Protein Sequence and Structure* Vol. 5, suppl. 3 (ed. Dayhoff, M. O.) 345–352 (National Biomedical Research Foundation, Silver Spring, Maryland, 1978).

ACKNOWLEDGEMENTS. We thank Drs J. Battey and K. Mearow for advice on PCR procedures and *in situ* hybridization, respectively. Drs J. Fex, P. Nelson and E. Freese for their support in carrying out this study and for reading the manuscript, and D. Schoenberg for her help in preparing this manuscript.

Molecular structure of the chick cerebellar kainate-binding subunit of a putative glutamate receptor

Paul Gregor, Itzhak Mano, Ilana Maoz,
Michael McKeown* & Vivian I. Teichberg

Department of Neurobiology, The Weizmann Institute of Science, Rehovot 76100, Israel

* Molecular Biology and Virology Laboratory, The Salk Institute, PO Box 85800, San Diego, California 92138, USA

KAINATE receptors mediate some of the excitatory transactions carried out in the central nervous system by the neurotransmitter glutamate¹. They are involved in neurotoxicity², possibly in neurodegenerative disorders³ and it has been suggested that they have a role in long-term potentiation⁴. Kainate receptors are present both on neuronal¹ and glial^{5–7} cell membranes where they regulate the gating of a voltage-independent ion channel⁸. Nothing is known about their molecular structure. Taking advantage of the unusually high abundance of ³H-kainate binding sites in the chick cerebellum^{9,10}, we have isolated an oligomeric protein that displays a pharmacological profile similar to that of a kainate receptor, and have demonstrated, using the monoclonal antibody IX-50, that this protein is composed of a single polypeptide of M_r 49,000 which harbours the specific kainate recognition site¹¹. The structure of this kainate binding protein (KBP) is also of interest because of its exclusive cerebellar localization on Bergmann glial membrane¹² in close proximity to established glutamatergic synapses¹³. We now report the isolation of the complementary DNA containing the complete coding region of the kainate binding protein. The predicted structure of the mature protein has four putative transmembrane domains with a topology analogous to that found in the superfamily of ligand-gated ion channels^{14–16}. This raises the possibility, that kainate binding protein may form part of an ion channel and may be a subunit of a kainate subtype of glutamate receptor.

The chick cerebellar kainate binding protein (KBP) was isolated as previously described¹¹, as a single polypeptide of M_r 49K with a purity, estimated on SDS-polyacrylamide gels, of 80–95%. KBPs with similar M_r values have been isolated from frog¹⁷ and pigeon¹⁸ brain. Microsequencing of the N terminus of purified KBP produced only one sequence. To obtain internal amino-acid sequences of KBP, the latter was fragmented with cyanogen bromide and the five most prominent peptides (I–V), obtained on separation by reverse-phase HPLC, were submitted to N-terminal sequence analysis (P.G. and V.I.T., manuscript in preparation). All sequences determined chemically are shown (solid lines) in Fig. 1.

As peptide IV was immunopositive with monoclonal antibody IX-50 and polyclonal antibodies (data not shown) we synthesized a set of 20-nucleotide-long oligodeoxyribonucleotides (oligo(dT)) corresponding to its N terminus (probe IV; see Fig. 2 legend). To establish its suitability for screening libraries and the optimal hybridization conditions, probe IV was used to analyse RNA blots. As shown in Fig. 2a (panel A), probe IV hybridized mainly to a 3.9-kilobase (kb) RNA, but also to a 6-kb RNA. Both of these are much more abundant in the cerebellum than in the cerebrum and not detectable in the liver. This tissue distribution correlates very well with that of kainate-binding sites¹⁹ and with anti-KBP immunoreactivity (data not shown). Analysis of RNA blots with probe I, derived from the sequence of peptide I, gives an identical pattern to the one shown in Fig. 2a, panel A.

An oligo(dT)-primed chick cerebellar cDNA library, constructed in phage λ gt-10, was screened with the probe IV (for methods, see Fig. 2 legend). Positive phages were rescreened with probe I and cDNAs KBP-2.9 and KBP-2.2 were isolated. Radiolabelled KBP-2.2 cross-hybridizes to KBP-2.9 and gives the anticipated hybridization pattern on northern blots (Fig. 2a, panel B). Rescreening the same library with restriction fragments of cDNA KBP-2.2 resulted in isolation of the clone KBP-3.3, which has the complete coding region of KBP. The maps of clones used to establish the coding DNA sequence of KBP are shown in Fig. 2b.

The entire protein-coding DNA sequence and deduced

amino-acid sequence of KBP are shown in Fig. 1. The mature protein is predicted to have 464 amino acids and an M_r of 51.8K. The 23 residues preceding the N-terminal sequence probably constitute the signal peptide. The first methionine is preceded by an in-frame stop codon and has a sequence context consistent with its use as a translation start site²⁰. Although we have not formally shown that the cloned cDNA expresses a protein with kainate-binding properties, several arguments indicate that the derived amino-acid sequence corresponds to that of KBP. First, the deduced amino-acid sequence contains all the peptide sequences derived from the pure 49K polypeptide harbouring the kainate-binding site¹¹. Second, two of the encoded peptides (IV and V) provide determinants for polyclonal antibodies directed against KBP and peptide IV is recognized by the anti-KBP monoclonal antibody IX-50. Third, the amino-acid composition and size of the deduced sequence are, within limits of error, consistent with those of isolated KBP. Finally, the preferential cerebellar localization of the RNA species hybridizing to the cloned cDNA is strictly in line with the selective tissue distribution of KBP.

Data-base searches did not reveal any significant sequence

homology of KBP to other proteins. A comparison of the overall pattern of hydrophobic residues for KBP with such patterns for known ligand-gated channels (Fig. 3), however, revealed some similarities. The KBP contains several hydrophobic regions including prominent ones near the N and C termini of the protein. In addition, there are three stretches of hydrophobicity between residues 140 and 240, each of which is long enough to span a lipid bilayer in an α -helical configuration. One interpretation of these data is that the first hydrophobic region functions as a signal peptide and that the other four regions (M1-M4) serve as membrane spanning regions. The size of KBP and the overall placement of the putative membrane spanning regions of KBP are similar to the size of known ligand-gated ion channel proteins and to the placement of putative membrane spanning regions of these proteins indicating that KBP has a general architecture similar to that of the chemically gated ion channels. If KBP resembles ligand-gated ion channels, the N and C termini are probably located extracellularly. This assignment is supported by the presence of putative N-glycosylation sites at positions 81 and 421, in the putative N- and C-terminal extracellular regions, and by the presence of putative phosphorylation sites

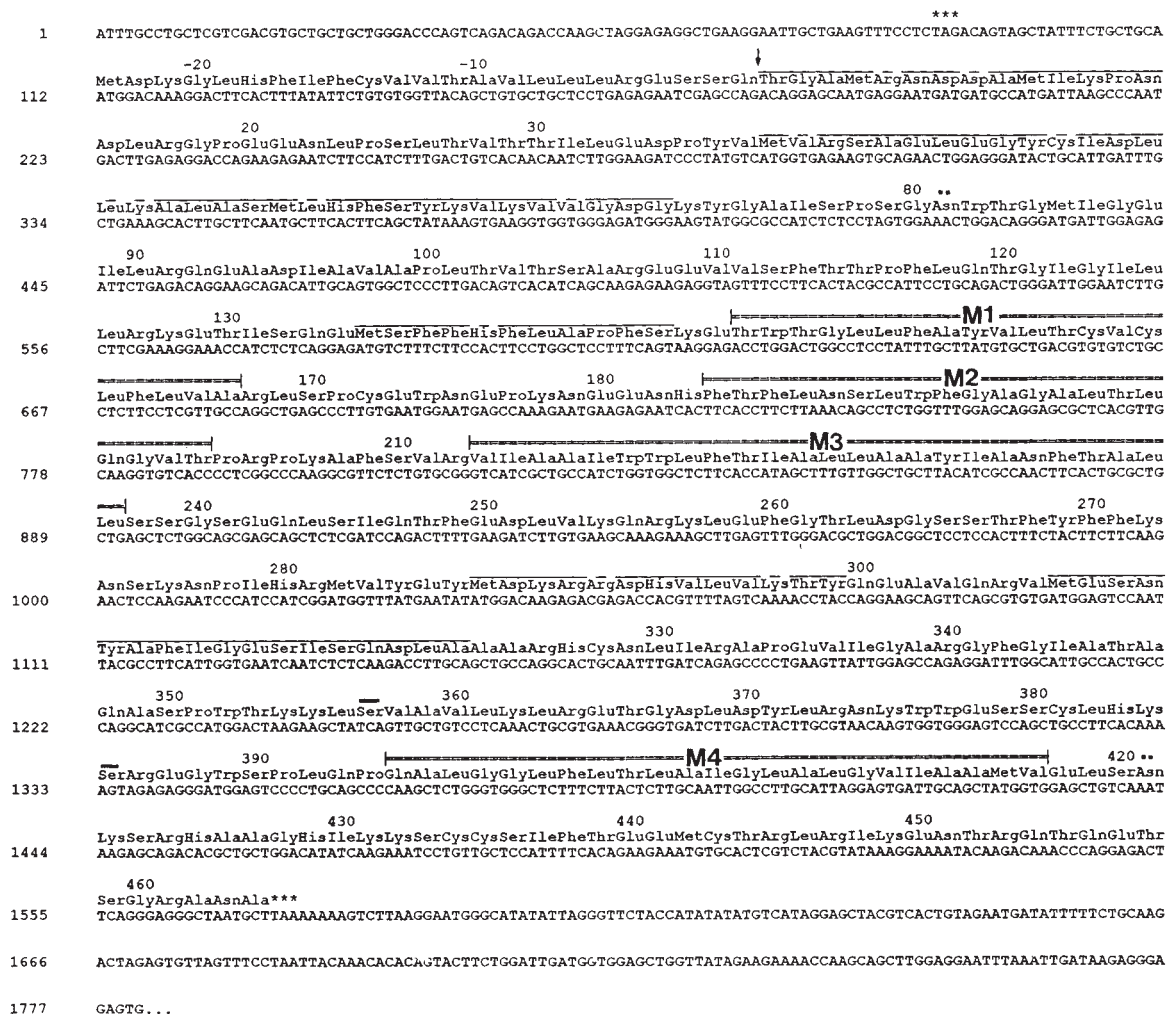


FIG. 1 DNA sequence and deduced amino-acid sequence of the chick cerebellar kainate binding protein. Nucleotides are numbered at left and amino acids above the sequence. The first in-frame stop codon in the 5' untranslated sequence and the stop codon at the end of the coding region are labelled with asterisks. Amino-acid numbering is from the first residue of the mature N-terminal sequence as established by direct N-terminal sequencing of the kainate binding protein. The presumptive signal peptide sequence is indicated by negative numbering. Peptide sequences determined chemically are denoted by solid lines and uncertain or unidentified residues by broken lines. Signal peptide cleavage site is indicated by arrow. The

proposed membrane-spanning hydrophobic sequences are indicated by double lines (M1-M4). The putative extracellular N-glycosylation sites are labelled with two dots and the putative intracellular phosphorylation sites by a solid bar.

METHODS. The cDNA inserts of positive phages were subcloned into Bluescript phagemid pKS vectors (Stratagene) and sequenced by the dideoxy chain termination method³⁰. The sequence presented is combined from all cDNAs and both strands were analysed. The 3' untranslated regions of the cDNA clones beyond the *Pst*I site were not sequenced completely on both strands and are not shown.

FIG. 2 Molecular cloning of the KBP cDNA. *a*, Northern blot analysis of chick RNA with oligodeoxyribonucleotide oligo-(dT) probe IV corresponding to peptide IV (panel A) and with cloned KBP-2.2 cDNA (panel B). Line 1, liver; line 2, cerebrum; line 3, cerebellum. *b*, cDNA clones and partial restriction map. All cDNAs have their cognate 3' poly-A sequences. The 3' untranslated region of KBP-2.2 beyond the *Hind* III site is different from KBP-3.3 and KBP-2.9, probably as a result of alternate 3'-end processing. The open reading frame of KBP-2.9 was truncated at its 5' side, probably as a result of incomplete splicing. Solid box, protein coding region, IVS, presumed intervening sequence.

METHODS. Standard recombinant DNA methods were as described²⁵. Based on the sequence of peptide IV, the oligo-(dT) 5'-AARGCGTAGTTGSWCTCCAT-3' (where R=A or G, S=C or G, W=A or T)²⁶ was synthesized (probe IV) and labelled to high specific radioactivity with [γ -³²P]ATP²⁷. Total RNA was isolated²⁸ from adult chick tissues; 25 μ g per line was electrophoresed on a 1% formaldehyde agarose gel, blotted on nitrocellulose membrane (Hybond-C, Amersham) and hybridized overnight with the radiolabelled oligo-(dT) probe at 37 °C in 5 $\times$ SSPE (1 $\times$ SSPE contained 0.18 M NaCl, 0.01 M sodium phosphate, 1 mM EDTA, pH 6.8) as described²⁷. The high stringency wash was at 37 °C in 0.5 $\times$ SSPE for 1 h. Antisense probe I based on peptide I sequence was synthesized as an oligonucleotide mixture 5'-TAIGTIACIACIACRTGRTICAIAC-QCTTRITC-3' (where I=inosine, Q=G or T and R=A or G), and radiolabelled as above. To isolate KBP cDNAs, an oligo-(dT)-primed cDNA library in phage λ gt-10 (ref. 29) was constructed from cerebellar poly(A)⁺-RNA of hatching chicks. Double-stranded cDNA was synthesized, fitted with *Eco*RI linkers, and cDNAs were size fractionated on 1% agarose gel. Fragments larger than 1.5 kb were isolated and ligated with *Eco*RI cut λ gt-10, resulting in 150,000 independent recombinants of which at least 50% carried inserts. The library was amplified and ~160,000 phages screened with probe IV as described above; positives were rescreened with probe I resulting in isolation of KBP-2.9 and KBP-2.2 clones. A screen of 300,000 phages, under high stringency conditions, with the 0.8 kb *Pst*I-*Pst*I and 0.1 kb *Eco*RI-*Pst*I fragments derived from the 5' end of KBP-2.2, led to the isolation of the clone KBP-3.3.

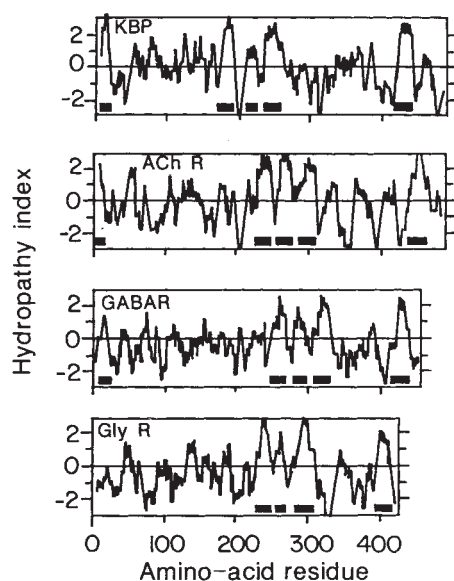
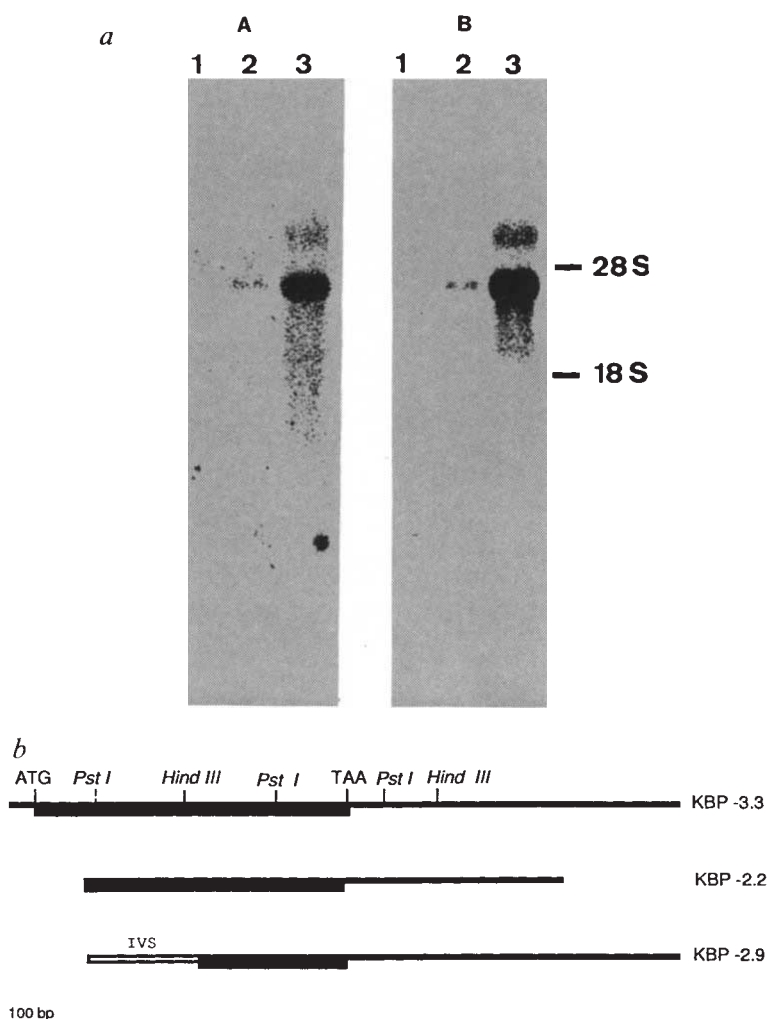


FIG. 3 Structural similarity of the KBP to subunits of ligand-gated ion channels. Hydropathy profiles computed according to Kyte and Doolittle³¹, averaged over a window of 9 residues and plotted with a 1-residue interval. Solid box, putative signal peptide and membrane spanning regions. AChR, α 2 chick acetylcholine receptor subunit³²; GABA R, γ -amino-butyric acid receptor α subunit¹⁵; Gly R, glycine receptor¹⁶.

at positions 357 and 385, in the putative intracellular region between M3 and M4.

Further analysis of KBP structure shows the presence of features compatible with channel function, particularly in its putative M2 transmembrane segment. Indeed, the M2 domain of the various nicotinic acetylcholine receptor (AChR) subunits¹⁴, which is essential to channel function²¹⁻²³, contains polar uncharged residues and is preceded by hydrophilic negatively charged residues. Similar features in the equivalent region of KBP are observed, indicating, only by analogy, a possible role in ion permeation. As mentioned above, KBP also contains potential phosphorylation sites, a finding consistent with the reported modulation of kainate-gated currents by protein kinase²⁴. If indeed, KBP forms part of an ion channel, its sequence does not display the homologies that have been used to group the various chemically gated ion channels in a single family. The structure of KBP must therefore reflect its idiosyncratic properties, one of which is its unusually high density in the cerebellum and exclusive presence on Bergmann glia (ref. 12 and P. Somogyi, N. Eshhar, V.I.T. & J. D. Roberts, unpublished observations). Although its function is not yet known, its localization on Bergmann glial membranes surrounding the parallel fibre—Purkinje cell glutamatergic synapse indicates a role in modulation of synaptic efficacy. Of relevance are our recent findings that isolated Bergmann glial cells in culture display high levels of KBP, depolarize in response to kainate and release GABA (γ -aminobutyric acid) (ref. 12 and A. Ortega and V.I.T., unpublished observations). The structural data presented here, together with the above observations, indicate therefore that KBP may be a subunit of a glial subtype of kainate receptor/cation channel. □

Received 18 October; accepted 14 November 1989.

1. Watkins, J. C. & Evans, R. H. A. *Rev. Pharmacol. Tox.* **21**, 165–204 (1981).
2. Fuxe, K., Roberts, P. & Schwarcz, R. (eds) *Excitotoxins* (Plenum, New York, 1983).
3. Cotman, C. W. & Monaghan, P. *Adv. exp. Med. Biol.* **203**, 237–252 (1986).
4. Collingridge, L. & Bliss, T. V. P. *Trends Neurosci.* **10**, 288–293 (1987).
5. Bowman, C. L. & Kimelberg, H. K. *Nature* **311**, 656–659 (1984).
6. Sontheimer, H., Kettenmann, H., Backus, K. H. & Schachner, M. *Glia* **1**, 328–336 (1988).
7. Usowicz, M. M., Gallo, V. & Cull-Candy, S. G. *Nature* **339**, 380–383 (1989).
8. O'Brien, R. & Fischbach, G. D. *J. Neurosci.* **6**, 3275–3283 (1986).
9. London, E. D., Klemm, N. & Coyle, J. T. *Brain Res.* **192**, 463–476 (1980).
10. Henke, H. & Cuenod, M. in *Neurotransmitters and Their Receptors* (eds Littauer, U. Z. et al.) 373–390 (Wiley, New York, 1980).
11. Gregor, P., Eshhar, N., Ortega, A. & Teichberg, V. I. *EMBO J.* **7**, 2673–2679 (1988).
12. Teichberg, V. I. et al. in *Excitatory Amino Acids and Neuronal Plasticity* (ed. Y. Ben Ari) (Plenum, New York, in the press).
13. Blackstone, C. D., Supattapone, S. & Snyder, S. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4316–4320 (1989).
14. Numa, S. *Harvey Lectures* **83**, 121–165 (1989).
15. Schofield, P. et al. *Nature* **328**, 221–227 (1987).
16. Grenningloh, G. et al. *Nature* **328**, 215–220 (1987).
17. Hampson, D. R. & Wenthold, R. J. *J. biol. Chem.* **263**, 2500–2505 (1988).
18. Klein, A. U., Niederost, B., Winterhalter, K. H., Cuenod, M. & Streif, P. *Neurosci. Lett.* **95**, 359–364 (1988).
19. Eshhar, N. et al. *Brain Res.* **476**, 57–70 (1989).
20. Kozak, M. *Nucleic Acids Res.* **15**, 8125–8132 (1987).
21. Imoto, K. et al. *Nature* **335**, 645–648 (1988).
22. Leonard, R. J., Labarca, C. G., Charnet, P., Davidson, N. & Lester, H. A. *Science* **242**, 1578–1581 (1988).
23. Miller, C. *Neuron* **2**, 1195–1205 (1989).
24. Lirman, E. R., Knapp, A. G. & Dowling, J. E. *Brain Res.* **481**, 399–402 (1989).
25. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1984).
26. Lathé, R. *J. molec. Biol.* **183**, 1–12 (1985).
27. Meinkoth, J. & Wahl, G. *Analyt. Biochem.* **138**, 267–284 (1984).
28. Cathala, G. et al. *DNA* **2**, 329–335 (1983).
29. Huynh, T. V., Young, R. A. & Davis, R. W. in *DNA Cloning: A Practical Approach* Vol. 1 (ed. Glover, D. M.) 49–78 (IRL, Oxford, 1985).
30. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463–5467 (1977).
31. Kyte, J. & Doolittle, R. F. *J. molec. Biol.* **157**, 105–132 (1982).
32. Nef, P., Oneyser, C., Alliod, Ch., Couturier, S. & Ballivet, M. *EMBO J.* **7**, 505–601 (1988).

ACKNOWLEDGEMENTS. We thank Natan Tal, David Ornstein and Chevi Lamed for technical assistance. This work was supported by the Leo and Julia Forchheimer Center of Molecular Genetics at the W.I.S. and by N.I.H. (M.M.K.).

Specificity pockets for the side chains of peptide antigens in HLA-Aw68

T. P. J. Garrett*, M. A. Saper*, P. J. Bjorkman†, J. L. Strominger & D. C. Wiley*‡

Department of Biochemistry and Molecular Biology and
*Howard Hughes Medical Institute, Harvard University, Cambridge,
Massachusetts 02138, USA

WE have determined the structure of a second human histocompatibility glycoprotein, HLA-Aw68, by X-ray crystallography and refined it to a resolution of 2.6 Å. Overall, the structure is extremely similar to that of HLA-A2 (refs 1, 2; and M.A.S. et al., manuscript in preparation), although the 11 amino-acid substitutions at polymorphic residues^{3,4} in the antigen-binding cleft² alter the detailed shape and electrostatic charge of that site. A prominent negatively charged pocket within the cleft extends underneath the α -helix of the α_1 -domain, providing a potential subsite for recognizing a positively charged side chain or peptide N terminus. Uninterpreted electron density, presumably representing an unknown 'antigen(s)', which seems to be different from that seen in the HLA-A2 structure¹, occupies the cleft and extends into the negatively charged pocket in HLA-Aw68. The structures of HLA-Aw68 and HLA-A2 demonstrate how polymorphism creates and alters subsites (pockets) positioned to bind peptide side chains, thereby suggesting the structural basis for allelic specificity in foreign antigen binding.

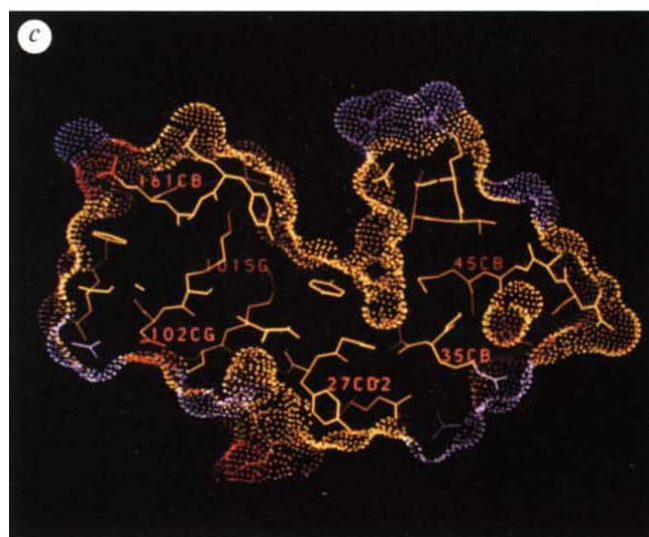
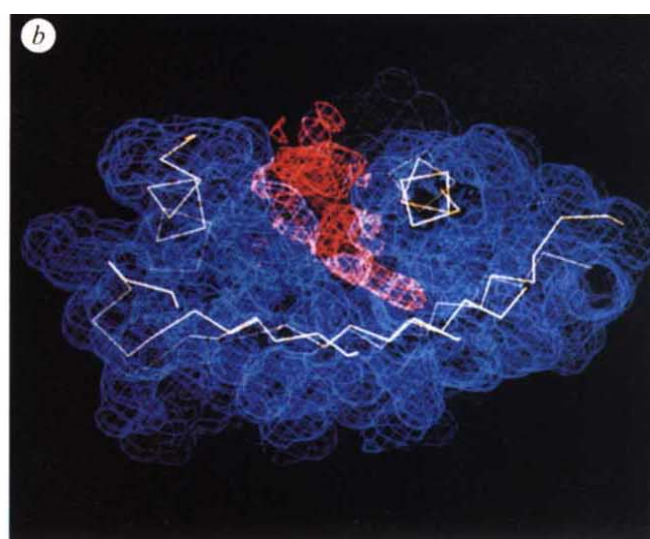
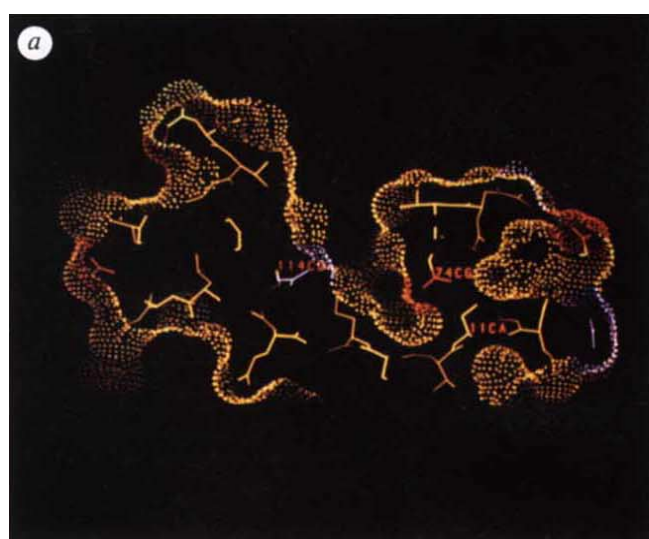
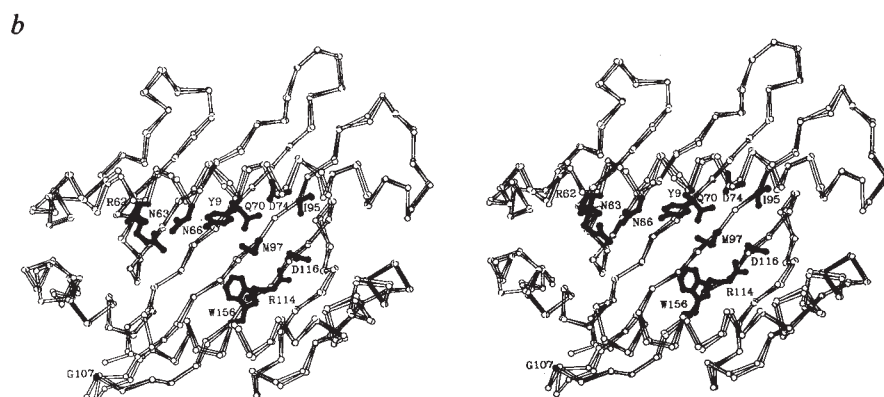
HLA-Aw68 (formerly HLA-A28) is the product of one allele of the polymorphic HLA-A locus of the major histocompatibility complex (MHC). Like other class I histocompatibility glycoproteins HLA-Aw68 seems to form complexes with peptides derived from internal degradation of self and nonself antigens^{5–8}. These complexes are recognized on cell surfaces by cytotoxic T lymphocytes (CTL) leading to the elimination of virally infected or histoincompatible cells. The extensive polymorphism of histocompatibility glycoproteins is concentrated at the putative peptide-binding cleft^{2,9}, indicating that the variability was selected to generate diversity in peptide binding and T-cell receptor recognition². Individual class I and class II histocompatibility glycoproteins interact with an extremely broad range of peptides of diverse sequences^{10–14}, yet exhibit selectivity by failing to interact with many peptides. This failure to interact probably

FIG. 1 Comparison of HLA-Aw68 and HLA-A2. *a*, Main-chain ribbon diagram of HLA-Aw68 with the 13 side chains of residues that differ from those of HLA-A2. Ten differences between HLA-A2 and HLA-Aw68, respectively, (F9Y, E63N, K63N, H70Q, H74D, V95I, R97M, H114R, Y116D and L156W) are in the binding cleft. One difference, G62R, faces up from the α_1 -domain α -helix (labelled 62). W107G (labelled 107) is on a loop in the α_2 -domain and A254V (labelled 245) is on the α_3 -domain. β_2 -microglobulin is red. *b*, Stereoview of HLA-Aw68 and HLA-A2 showing the α -carbon atoms of the α_1 - and α_2 -domains superimposed. Side-chain differences in HLA-Aw68 are in solid lines. Root-mean-square differences between atomic positions of α -carbons (with $B \leq 35 \text{ Å}^2$) in HLA-Aw68 and HLA-A2: domains α_1 and α_2 , 0.45 Å; α_3 -domains, 0.27 Å; and β_2 -microglobulin, 0.30 Å; Standard deviation (s.d.) of overall coordinate errors in each structure are estimated at 0.3 Å (for further details, see refs 18, 19, and M.A.S. et al., manuscript in preparation). (In the site only a few small shifts in main-chain positions (over 2 s.d. of the error) are observed, one in the α -helix of α_1 -domain at residues 61–71, another in the α -helix of α_2 -domain in the region 148–156, and a third in a β -strand at positions 97, 98 and 101 (Fig. 1*b*)). For details of structure determination, see ref. 18. Briefly, the initial model was based on the structure of HLA-A2 at 3.0 Å. Side-chains that differed between the two molecules were fit in electron density maps calculated with coefficients F_o (Aw68) and F_o (Aw68)– F_o (A2) and phases calculated from the HLA-A2 model. The structure was refined using the program, XPLOR^{30,31}; current R factor = 17.9% for all data; $I > 1\sigma$; 6–2.6 Å; 7 water molecules; r.m.s. deviations from ideal bond lengths, 0.016 Å; angles, 3.5°. When residues that differ between HLA-Aw68 and HLA-A2 were omitted, difference maps showed clear electron density for the ten residues pointing into the site and for 245. Density for Gly 107 and the side-chain of Arg 62 appeared weak owing to disorder. An analysis of the effect of crystal contacts on the structures of HLA is found in ref. 17. *a*, Prepared with HYDRA (R. Hubbard, unpublished); *b*, Prepared with PLUTO (S. Motherwell, E. Dodson and P. Evans, unpublished). Single-letter amino-acid code.

FIG. 4 Sections perpendicular to the length of the cleft showing different and similar pockets in HLA-Aw68 and HLA-A2 binding sites. *a*, *c*, *d*, Sections through solvent contact surfaces coloured by charge: red, acidic; blue, basic; green, His; yellow, neutral. The cleft appears as a central depression; the α_1 -domain α -helix is towards the top right, the α_2 -domain α -helix is at the top left; β -sheet residues form the bottom of the groove. *a*, A pocket at Asp 74 under the α_1 -domain α -helix of HLA-Aw68. Note red 'dots' on the top of the pocket indicate the surface is formed by the carboxylate oxygens of Asp, 74. *b*, Same view of HLA-Aw68 as in *a* showing extra electron density (red) which fills the cleft, reaching into the '74 pocket'. Picture is sectioned through a blue van der Waal's surface of HLA-Aw68 and has a yellow α -carbon backbone. *c*, The pocket at Met 45 of HLA-Aw68 extends under the α_1 -domain α -helix. *d*, The pocket at Met 45 HLA-A2 is similar to that in *c*. Note also the essentially identical structure of HLA-Aw68 and HLA-A2 in regions where atoms are labelled in red. (Label: Sequence number, atom type). The extra electron density bound to HLA-Aw68 is not currently interpretable as a single conformation such as an α -helix. Figures constructed using HYDRA (R. Hubbard, unpublished) and MS (ref. 29).

† Present address: Division of Biology, Howard Hughes Medical Institute, California Institute of Technology, Pasadena, California, USA.

‡ To whom correspondence should be addressed.



provides the basis for allele-specific nonresponsiveness to certain immunological challenges^{10,11,15,16}. Here we report the structure of HLA-Aw68 refined to a resolution of 2.6 Å, which when compared with the 2.7 Å resolution structure of HLA-A2 (M.A.S. *et al.*, manuscript in preparation, and ref. 17) suggests mechanisms for both the broad specificity and the allelic selectivity of peptide binding.

The structure of papain-solubilized HLA-Aw68 was determined and refined to 2.6-Å resolution (see ref. 18 for details) from orthorhombic crystals isomorphous to HLA-A2 crystals¹⁹.

The overall structure, excluding the 13 amino-acid differences, is very similar to that of HLA-A2 refined to high resolution from both monoclinic crystals (2.7 Å; M.A.S. *et al.*, manuscript in preparation) and orthorhombic (2.5 Å) crystals¹⁷. Figure 1b shows the α -carbon backbone of HLA-Aw68 overlaid on the coordinates of the 2.7 Å resolution monoclinic crystal structure of HLA-A2 and lists root mean square (r.m.s.) deviations in coordinates between domains of the two structures. The average deviations are within the expected experimental error in the coordinates of these structures (Fig. 1b, legend).

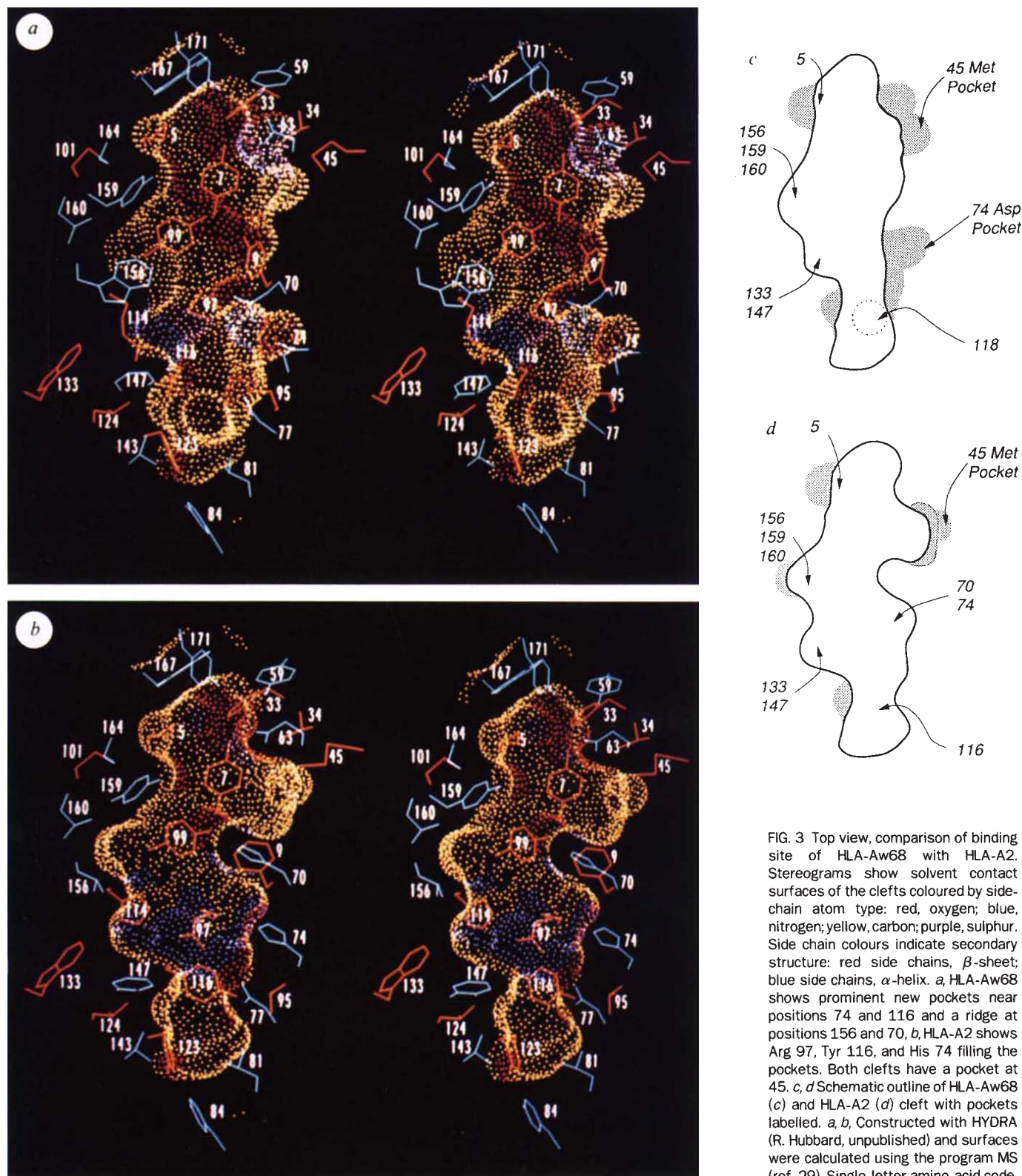
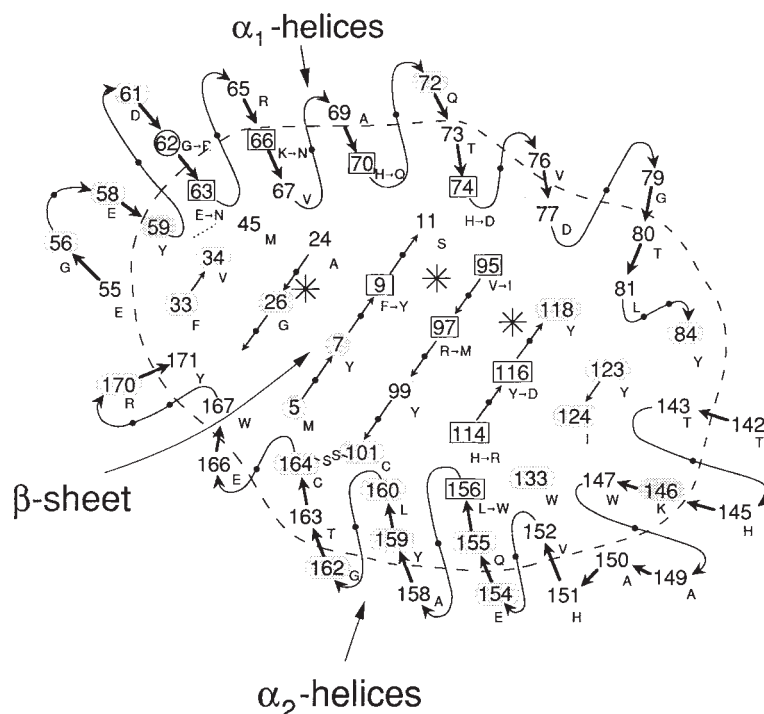


FIG. 3 Top view, comparison of binding site of HLA-Aw68 with HLA-A2. Stereograms show solvent contact surfaces of the clefts coloured by side-chain atom type: red, oxygen; blue, nitrogen; yellow, carbon; purple, sulphur. Side chain colours indicate secondary structure: red side chains, β -sheet; blue side chains, α -helix. a, HLA-Aw68 shows prominent new pockets near positions 74 and 116 and a ridge at positions 156 and 160 and Arg 97, Tyr 116, and His 74 filling the pockets. Both clefts have a pocket at 45. c, d Schematic outline of HLA-Aw68 (c) and HLA-A2 (d) cleft with pockets labelled. a, b, Constructed with HYDRA (R. Hubbard, unpublished) and surfaces were calculated using the program MS (ref. 29). Single-letter amino-acid code.

FIG. 2 Diagram of class I binding cleft. Thin arrows (at centre) are β -sheet residues on the cleft floor. Helical residues are shown at both edges, bold arrows are the top helical surface. Sequence labels are HLA-A2, with HLA-Aw68 shown after a short arrow (62G $\rightarrow$ R). Residues inside dashed line are ≤ 4.5 Å from extra electron density and some may contact peptides. Differences between HLA-A2 and HLA-Aw68 are boxed. Stippled residues are conserved in all human class I sequences^{9,28}. Asterisks mark pockets described in text. Single-letter amino-acid code.



The principal differences between the structures of HLA-Aw68 and HLA-A2 are in the size and precise location of the side chains of the 13 amino-acid differences (Fig. 1). Ten of these face into the putative peptide-binding cleft: residue 62 points 'up' from the α_1 -domain α -helix in a position where it could be recognized directly by a T-cell receptor²; residue 107 is on a loop between β -strands of the α_2 -domain; and residue 245 is on the α_3 -domain at the putative contact site with the accessory molecule CD8 (refs 20, 21). The number of amino acids that differ within the site (10) is approximately equal to the average number of site differences between pairs of sequences within HLA loci (9.7). Therefore the HLA-Aw68-HLA-A2 comparison should provide a representative view of the effect of polymorphic changes within the antigen-binding site. Figure 2 summarizes diagrammatically the location of the 10 differences in the antigen-binding cleft and how they affect the lining of the cleft. For example, the asterisks indicate pockets formed in the checkerboard pattern of β -sheet side-chains between residues 9, 11, 97, 95 and 74 and between residues 95, 97, 116 and 118 that are opened up in HLA-Aw68 by the substitutions at residues 74, 95, 97, 114, and 116.

Although each individual amino-acid difference in the binding cleft seems to cause only very local structural changes (limited to a few tenths of angstroms in nearest neighbour side chains), the 10 substitutions taken together substantially transform the shape and charge of parts of the cleft but leave other parts unchanged (Fig. 3). Figure 4a illustrates the most dramatic difference between the two structures—a deep negatively charged pocket formed below the α_1 -domain α -helix of HLA-Aw68 by the substitution of Asp for His at position 74. Access to this pocket is facilitated by the substitutions of Met for Arg at position 97 and Asp for Tyr at position 116 on the α_2 -domain β -sheet (compare Fig. 3a and b). Figure 4b shows that the 'extra electron density' which fills the antigen-binding cleft extends into this pocket. The top surface of the pocket is formed by the carboxylate oxygen atoms of Asp 74 (red dots, Fig. 4a), so that this pocket seems suited to bind polar atoms, especially a positively charged side-chain or N terminus of an antigenic peptide. The sequences of other class I glycoproteins indicate that a pocket with this specificity is probably present in several

HLA-A and -B molecules. Other prominent alterations in HLA-Aw68 include a deep pocket in the α_2 -domain β -sheet caused by the substitution of the small amino acid, aspartic acid, for a large aromatic tyrosine at position 116 (see Fig. 3a; asterisks in Fig. 2), and the presence of a ridge across the centre of the bottom of the cleft owing to the bulky substitutions of tryptophan at position 156 and glutamine at position 70 (Fig. 3a).

One prominent pocket below the α_1 -domain α -helix remains essentially conserved between the two structures (Fig. 4c, d). This pocket is a hydrophobic cavity with Met 45 at its base (surrounded by Val 67, Val 34, Gly 26 and Ala 24; asterisks in Fig. 2, and Fig. 3). It seems large enough to bind a leucine side chain (M.A.S. *et al.*, in preparation). Comparison of HLA sequences at residues in and around this pocket indicate that it is present in many HLAs including HLA-B27 and HLA-B44. In HLA-B27, position 45 is a negatively charged glutamic acid, and in HLA-B44 it is a positively charged lysine. Thus the '45 pocket' is formally reminiscent of the specificity site in serine proteases, a pocket that is hydrophobic in chymotrypsin but negatively charged in trypsin. In HLA-peptide complexes it is possible that this pocket binds nonpolar side chains in HLA-Aw68, and positively and negatively charged polar side chains in HLA-B27 and HLA-B44, respectively. Large substitutions at position 67 at the entry of the pocket could close it in some allelic products such as HLA-B7 (Tyr at position 67). Most of the residues lining the pocket and around its opening are polymorphic, thereby possibly changing the pocket specificity in different HLA alleles. Taurog and El-Zaatari²² have presented evidence that raises the possibility that a free cysteine at position 67 in HLA-B27 is covalently modified in the HLA-B27-linked disease ankylosing spondylitis. Given the location of position 67 at the edge of a negatively charged pocket in HLA-B27, it might be possible to design, perhaps from an antigenic peptide, a covalent blocking agent incorporating a positive charge to complement the '45 pocket' and a thiol to react with cysteine 67. Consideration of pockets might be sufficient to predict peptide binding even when alleles differ by a substantial number of residues. The sequence of HLA-Aw69, which is a natural hybrid of α_2 - and α_3 -domains from HLA-A2 and the α_1 -domain from HLA-Aw68 (ref. 3), indicates that the '45 pocket' would

be present but that the Asp 74 pocket would probably be blocked by Arg 97 and Tyr 116 (as in HLA-A2), possibly explaining why one HLA-A2-specific peptide antigen also interacts with HLA-Aw69 (ref. 23).

Comparison of the three-dimensional structures of HLA-Aw68 and HLA-A2 demonstrates how different pockets or subsites within the cleft of the class I histocompatibility glycoproteins are produced by substitutions at positions of amino-acid polymorphism (compare Fig. 3c and d). The difference between these two allele products is typical of MHC allelic diversity, and so the same degree of structural change in pockets and subsites should be observed for other alleles. In the case we have described, polymorphism results in limited local structural changes; the backbone structure is generally not altered (Fig. 1b), and residues identical in both structures remain essentially unperturbed¹⁸. The observed effects of polymorphism

were on the shape (for example, intrusions and ridges) and charge-distribution of the antigen-binding site, and on some local pockets, resulting in changes in their specificity and in their distribution in the site. These structural changes provide the basis for allelic specificity in peptide affinity^{11,15,16} and the resultant non-responsiveness (and possibly hyper-responsiveness) to certain immunological challenges¹⁰. Because amino-acid side chains from antigenic peptides could bind in these pockets, peptides binding to one MHC allelic product could then have a particular pattern of side chains^{24,25} to fit the pockets. Moreover, because the affinity of antigenic peptides for individual histocompatibility glycoproteins may depend on only a few of an antigenic peptide's side-chains fitting into a subset of the available pockets^{26,27}, a broad range of peptides could be accommodated by the various combinations of filled and empty pockets. □

Received 18 August; accepted 12 October 1989

1. Bjorkman, P. J. *et al. Nature* **329**, 506–512 (1987).
2. Bjorkman, P. J. *et al. Nature* **329**, 512–518 (1987).
3. Holmes, N. & Parham, P. *EMBO J.* **4**, 2849–2854 (1985).
4. Koller, B. H. & Orr, H. T. *J. Immunol.* **134**, 2727–2733 (1985).
5. Townsend, A. R. M., Gotch, F. M. & Davey, J. (1985) *Cell* **42**, 457 (1985).
6. Germain, R. M. *Nature* **322**, 687 (1986).
7. Bevan, M. *Nature* **325**, 192 (1987).
8. Townsend, A. & Bodmer, H. A. *Rev. Immunol.* **7**, 601–624 (1989).
9. Parham, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 4005–4009 (1988).
10. Babbitt, B., Allen, P., Matsueda, E., Haber, E. & Unanue, E. *Nature* **317**, 359–361 (1985).
11. Buus, S., Sette, A., Colon, S. M., Miles, C. & Grey, H. M. *Science* **235**, 1353–1358 (1987).
12. Sette, A., Buus, S., Colon, S. M., Miles, C. & Grey, H. M. *J. Immunol.* **141**, 45–48 (1988).
13. Townsend, A. R. M. *et al., Cell* **44**, 959–968 (1986).
14. Maryanski, J. L., Pala, P., Coradin, J., Jordan, B. B. & Cerotini, J. C. *Nature* **324**, 578–579 (1986).
15. McDevitt, H. O. & Tyan, M. L. *J. expl. Med.* **128**, 1–11 (1968).
16. Fox, B. S., Carbone, F. R., Germain, R. N., Paterson, Y. & Schwartz, R. H. *Nature* **331**, 538–540 (1988).
17. Madden, D. R. *et al. Cold Spring Harb. Symp. Quant. Biol.* **54**, (in the press).
18. Garrett, T. P. J., Saper, M. A. & Wiley, D. C. *J. molec. Biol.* (submitted).
19. Bjorkman, P. J., Strominger, J. L. & Wiley, D. C. *J. molec. Biol.* **186**, 205–210 (1985).

20. Salter, R. D. *et al. Nature* **338**, 345–347 (1989).
21. Connolly, J. M., Potter, T. A., Wormstall, E.-M. & Hansen, T. H. *J. exp. Med.* **168**, 325–341 (1988).
22. Taurog, J. D. & El-Zaatari, A. K. *J. clin. Invest.* **82**, 987–992 (1988).
23. Bodmer, H. C., Gotch, F. M. & McMichael, A. J. *Nature* **337**, 653–655 (1989).
24. Sette, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 3296–3300 (1989).
25. Rothbard, J. B. & Taylor, W. R. *EMBO J.* **7**, 93–100 (1988).
26. Jardetzky, T. S. *et al. EMBO J.* (submitted).
27. Maryanski, J. L., Verdini, A. S., Weber, P. C., Salemme, F. R. & Coradin, G. *Cell* (submitted).
28. Ellis, S. A., Strachan, T., Palmer, M. S. & McMichael, A. J. *J. Immunol.* **142**, 3281–3285 (1989).
29. Connolly, M. L. *Quantum Chemistry Program Exchange Bull.* **1**, 74.
30. Brünger, A. T., Kuriyan, J. & Karplus, M. *Science* **235**, 458–460 (1987).
31. Brünger, A. T. in *Topics in Molecular Biology* (ed. Goodfellow, J. M.) (in the press).

ACKNOWLEDGEMENTS. We thank Anastasia Haykov for technical assistance, Ted Jardetzky and Peter Parham for comments on the manuscript, and acknowledge support from the NIH and Howard Hughes Medical Institute. We also thank Drs Dean Mann (NIH), D. Michael Strong and James Wood (US Naval Research Unit) and Don Giard (MIT Cell Culture Facility) for provision of cells. Coordinates of HLA-Aw68 have been deposited in the Brookhaven Protein Data Bank (Accession no. 2HLA).

Limited receptor repertoire in a mycobacteria-reactive subset of $\gamma\delta$ T lymphocytes

Mary Pat Happ*, Ralph T. Kubo†‡, Ed Palmer*‡, Willi K. Born*†‡ & Rebecca L. O'Brien†§

*Department of Pediatrics, Division of Basic Sciences and †Department of Medicine, National Jewish Center for Immunology and Respiratory Medicine, Denver, Colorado 80206, USA

‡Department of Microbiology and Immunology, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

§Howard Hughes Medical Institute, Denver, Colorado 80206, USA

THE physiological role of lymphocytes bearing the $\gamma\delta$ T-cell receptor (TCR) is still unclear. A function for a subset of these cells, however, is inferred from the finding that certain $\gamma\delta$ chain-bearing lymphocytes are stimulated in a receptor-dependent fashion by mycobacterial antigens^{1–5}. We found that hybridomas derived from such cells in newborn murine thymus not only responded to mycobacterial purified protein derivative (PPD), but also exhibited an apparent autoreactivity¹. In neither response was haplotype-specific major histocompatibility (MHC) restriction demonstrable. To investigate the nature of antigen recognition by these $\gamma\delta^+$ cells, we sequenced the γ - and δ -chains from 28 PPD-reactive hybridomas, and found that a specific γ -chain, together with one of a limited set of δ -chains, was needed to generate the PPD specificity. The reactive $\gamma\delta$ pairs exhibited considerable junctional diversity, which may act to produce differences in the fine specificities of the responding cells.

The TCR proteins from several PPD-reactive $\gamma\delta$ T-cell hybridomas¹ were compared with those from a collection of non-PPD-reactive $\gamma\delta$ -hybridomas¹ by SDS-PAGE (Fig. 1, and data not shown). Most of the PPD-reactive hybridomas shared

a δ -chain (the upper band of the receptor doublet) that was not frequently found among the nonreactive population (see also ref. 1). Although γ -chains showed a wider variability in relative molecular mass (M_r) than did δ -chains, all of the PPD-reactive and two of the nonreactive hybridomas possessed a γ -chain of the same M_r (39,000; the lower band of the receptor doublet). The size of this γ -chain indicated that it contained C γ 4 segment (ref. 6). Using a V γ 1–C γ 4 combination of oligonucleotide primers, polymerase chain reaction (PCR) amplification products were generated from the messenger RNA of all 28 of the PPD-reactive hybridomas and from 7 of the 25 non-PPD-reactive hybridomas. Sequence analysis revealed an in-frame V γ 1JC γ 4 transcript⁷ (Fig. 2a) in all of the PPD-reactive hybridomas and only two of the non-PPD-reactive hybridomas. There did not seem to be selection for a particular VJ junctional sequence among the PPD-reactive hybridomas (Fig. 2b).

Based on our previous analysis, δ -chain gene oligonucleotides were designed that would amplify the V δ 6.1 gene segment (V δ m23 in refs 1 and 8), as well as the closely related V δ 7 and V δ 6 gene family members^{8–12}. Amplification products were generated with these oligonucleotides from 25 of the PPD-reactive and 3 of the non-PPD-reactive hybridomas. Sequence analysis revealed the use of two different V region gene segments (Fig. 2a). One was identical to the previously described V δ 7.1 (ref. 10). We have designated this V δ gene segment V δ 6.3. The partial sequence obtained from the other V gene segment was identical to that previously described for the plasmid p λ 12 member of the V δ 6 family¹³. The predicted junctional amino acids for all of the V δ 6 chains are listed in Fig. 2b, with the three non-PPD-reactive hybridomas indicated by asterisks. All of the V δ 6.3 and V δ 6 (p λ 12)-expressing hybridomas used the D δ 2 and J δ 1 gene segments, as do most $\gamma\delta$ T cells^{8,11,14,15}. Characteristic of adult (as opposed to fetal) $\gamma\delta$ cells^{11,14}, D δ 1 sequences were present in two of the hybridomas (BNT-9.12 and BNT-21.32), each in a different reading frame. This, and the presence of substantial

N region diversity in most of the PPD-reactive hybridomas, indicate that these cells appear relatively late in ontogeny¹¹. All three different reading frames of the *Dδ2* gene segment could be used with retention of PPD specificity. Also, expression of different *N* region-encoded amino acids did not abolish the reactivity.

In both the δ - and the γ -chain, there were no apparent differences between the junctional regions in PPD-reactive versus non-PPD-reactive hybridomas that could account for the difference in reactivity. Instead, PPD-reactivity seems to result from the pairing of a $V\gamma1JC\gamma4$ chain with a *Vδ6*-bearing chain. Twenty-five of the twenty-six hybridomas having this in-frame *V* gene combination exhibited PPD-reactivity; it is not yet clear why the remaining hybridoma (BNT-11.54) was nonreactive. Thus, whereas specificity for both the antigen and the presenting molecule seems to be encoded by the germline *V* gene segments, the diversity seen in the VD and VDJ joints of reactive hybridomas could serve simply to fine tune the response.

We next attempted to identify the *Vδ* gene segments used in the three *Vδ6*⁺ but PPD-reactive hybridomas (BNT-21.2, 41BNT-24 and 41BNT-117). Based on previous hybridization data¹, we performed PCR analysis with a *Vα2* oligonucleotide. We found that 41BNT-117, as well as two of the non-PPD-reactive hybridomas, possessed an in-frame transcript containing a new *Vδ* gene, which we have designated *Vδ8*. Both the nucleotide sequence and predicted corresponding amino-acid sequence of *Vδ8* (Fig. 3a and b, respectively) are >90% identical to those of members of the *Vα2* gene family. The *Vδ8*⁺ PPD-reactive hybridoma expressed an in-frame $V\gamma1JC\gamma4$ transcript, whereas the two non-reactive *Vδ8*⁺ hybridomas did not.

A summary of *Vγ* and *Vδ* gene segment use in the 28 PPD-

reactive and 25 non-reactive hybridomas is given in Table 1. The absolute requirement for the $V\gamma1JC\gamma4$ chain in the PPD-reactive hybridomas is conspicuous. Note that the *Vγ2* gene, although nearly 90% identical to *Vγ1* at the predicted amino-acid level⁷, was never found to encode the γ -chain expressed

a

	10	20	30	40
<i>Vγ1</i>	QLEQTELSVTRATDESAQISCIVSLPYFSNTAIHWYRQAKKFEYLIVVS			
31 clones	-----	-----	-----	-----
<i>Vα7.1</i>	EKVIQVWSTASRQEGEELTLDSCSYETSQVLYHLFWYKHLISGEMVFLIR			
<i>Vδ6.3</i> (23 clones)	-----	-----	-----	-----
<i>Vδ6(pλ12)</i>	QR-T--QP--S-W--V-----EYF-RI---RQ-F-----Y			
5 clones	W--V-----EY			

	50	60	70	80	90
<i>Vγ1</i>	TNYNQRPLGGKNNKIEASKDFQTSTLTKINYLKKKEDEATYYCAVW				
31 clones	-----	-----	-----	-----	-----
<i>Vα7.1</i>	QTSSSTAKERSGRYSVVFQKSLKSLISLIISALQPDSDGKYFCALWE				
<i>Vδ6.3</i> (23 clones)	-----	-----	-----	-----	-----
<i>Vδ6(pλ12)</i>	-P-FD-QNQ-----F-----V--S--E--T-----S-				
5 clones	--S--E--T-----S-				

b

Clone	<i>Vδ</i>	N	<i>Dδ1</i>	<i>Dδ2</i>	N	<i>Jδ1</i>	<i>Vγ1</i>	N	<i>Jγ4</i>
9.12	ALWE	L	WH	IGGIR	A	TDKLV	YCAVW		SGTSW
23.4	----	H		-----	-	TDKLV	-----		-----
10.7	----	H		-----	-	-----	-----	T	-----
9.4;23.2	----	H		-----	-	-----	-----	I	-----
40BNS-8	----	A		-----	-	-----	-----		-----
13.6	----	R		-----	-	-----	-----	I	-----
23.16	----	-V		-----	-	-----	-----		-----
3.12;18.13;23.12	----	-V		-----	-	-----	-----	I	-----
4.1	----	-V		-----	-	-----	-----		-----
13.17;15.4;15.15	----			-----	-	-----	-----	I	-----
8.29	----	-		-----	-	-----	-----		-----
*18.8	----	-		-----	-	-----	(not transcribed)		-----
10.16	--S	N		-----	-	-----	-----		-----
21.26	--S	-		-----	-	-----	-----	IT	-----
*13.13	--S	-		-----	-	-----	(non-functional)		-----
18.17	--S	-		-----	-	-----	-----	I	-----
19.8	----	-Y		RRDT	T	-----	-----	IA	-----
11.6	----	-D		RRDT	T	-----	-----		-----
18.5	----	-D		RR	D	-----	-----	IR	-----
*11.54	----	-AG		DT	-	-----	-----	Y	-----
14.5	----	H		R	SA	-----	-----	I	-----
21.32	--S	KD	GIY	RR	D	-----	-----	Y	-----
23.21	----	-		SEG	Y	-----	-----	IS	-----

FIG. 2 a, Predicted amino-acid sequences encoded by the *Vγ1* and *Vδ6* gene segments (single-letter amino-acid code). Nucleotide sequences are identical to those previously published for *Vγ1* (ref. 6), *Vα7.1* (ref. 10) and *Vδ6.3* (ref. 12). *Vδ6.3* and *Vδ6(pλ12)* have 69% amino-acid identity; the *Vδ6.3* gene segment is more closely related to the *Vδ6.1* (refs 8, 12; 86% amino-acid identity) and *Vδ6.2* (ref. 12; 97% amino-acid identity) family members. b, Predicted amino-acid sequences of the V-J junctional regions of the δ - and γ -chain mRNA found in the 28 *Vδ6*⁺ hybridomas. Hybridomas are from the 'BNT' collection (ref. 27; for example, BNT-9.12) unless otherwise noted. Hybridomas marked with an asterisk do not demonstrate PPD reactivity. No γ -chain junctional sequence is shown for those hybridomas which do not possess a functional $V\gamma1JC\gamma4$ transcript (BNT-18.8 and BNT-13.13). This γ -chain gene is not rearranged in the BW fusion line¹⁴. All clones bearing identical amino-acid sequences in the junctions were identical at the nucleotide level as well.

METHODS. RNA was isolated as previously described³¹. PCR amplification and sequence analysis of the amplified products was performed as previously described³² using the following oligonucleotide primers (5' to 3'): 5' *Vα7-Vδ6* external, GGATCTAATGTGGCCGAGA or GGATCTAATGTGGCCGAGA (amino-acid position, -4 to +3); 5' *Vα7* internal, AAGTGATTCAGGTCTGGTCA (amino-acid position, +3 to +9); 5' *Vδ6* internal, AGAAAGTGACTCAGGTCCAG (amino-acid position, +2 to +8); 5' *Cδ* external (antisense), ACAAGCAACATTTGTTCAT (amino-acid position, +8 to +15 of *Cδ*); 5' *Cδ* internal (antisense), TTTTCATGATGAACAGAT (amino-acid position, +15 to +21 of *Cδ*); 5' *Vγ1* external, GGGCTTGGGCAGCTGGAGCA (amino-acid position, -3 to +4); 5' *Vγ1* internal, AGTATCTAATATATGTCTCA (amino-acid position, +1 to +7); 5' *Cγ4* external (antisense), GAAGGAAG-GAAAATAGTAGG (amino-acid position, +12 to +18 of *Cγ*); 5' *Cγ4* internal (antisense), GGAGAAAAGTCTGAGTCAGT (amino-acid position, +4 to +10 of *Cγ*).

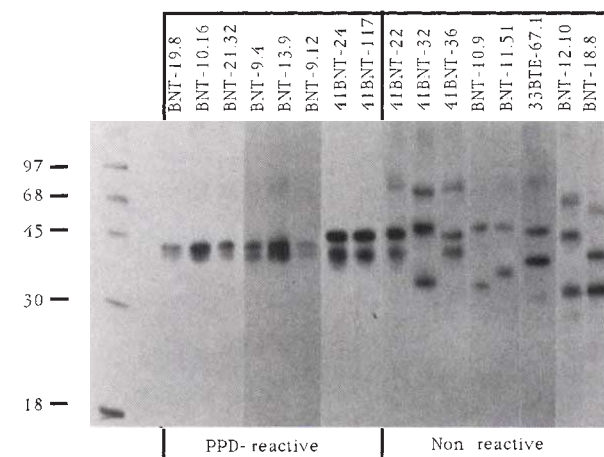
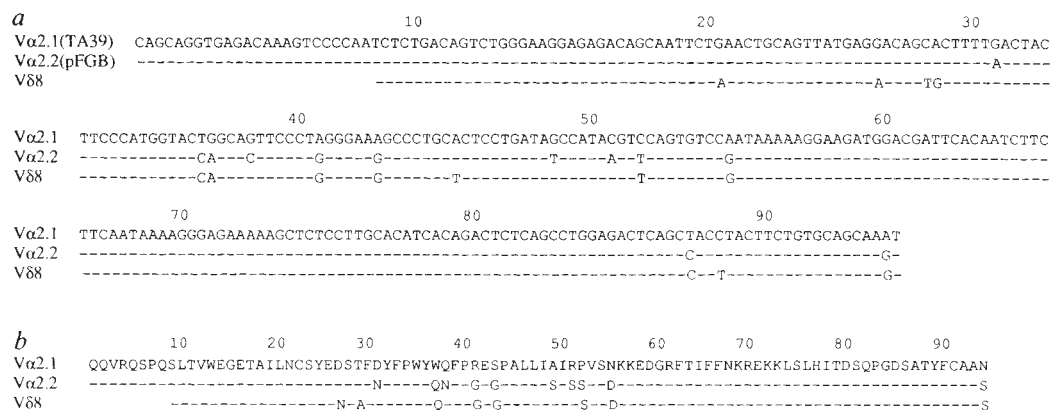


FIG. 1 Examination of surface TCR on $\gamma\delta$ hybridomas by SDS-PAGE under reducing conditions. Hybridomas were generated and tested for antigen reactivity as previously described^{1,27}. Representatives from both the PPD-reactive and the non-reactive collections are grouped accordingly. Bands corresponding to *M_s* of >50,000 (50K) are residual nonreduced heterodimer.

METHODS. Cell-surface radioiodination of T-cell hybridomas was performed using Iodogen (Pierce, Rockville, Illinois) as described previously²⁸. Cell lysates were prepared by detergent lysis using 1% digitonin²⁹. Immunoprecipitations of the cell lysates were performed as described previously using the anti-CD3 monoclonal antibody HMT3-1 (R.T.K. *et al.*, manuscript in preparation) with protein A-Sepharose (Sigma)³⁰. The immunoprecipitates were solubilized by boiling in Laemmli sample buffer³⁰ containing 0.01 M iodoacetamide, and electrophoresed on a SDS-PAGE gel (10% polyacrylamide). After electrophoresis, the gel was stained, destained, dried and autoradiographed for 18 h. Using the autoradiograph to localize the $\gamma\delta$ heterodimer, each $\gamma\delta$ band was excised from the dried gel and rehydrated in Laemmli sample buffer containing 5% 2-mercaptoethanol for 1-2 h. The gel slices were loaded onto a second SDS-PAGE gel (10% polyacrylamide) and electrophoresed to resolve the γ - and δ -subunits. The autoradiographs were developed after 3-5 days exposure. The positions of *M_r* protein standards (*M_r* × 10⁻³) are indicated on the left.

FIG. 3 Nucleotide sequences (a) and corresponding predicted amino-acid sequences (b) of *Va2.1* (isolated from C57BL/Ka (ref. 10) but identical to C57BL/10 (M.P.H., D. DiGiusto and E.P., unpublished observation)), *Va2.2* (isolated from BALB/c mice; M. Zauderer, personal communication) and *Vδ8* (designated *Vδ7* in ref. 1 just before a different *Vδ7* sequence was described in ref. 14; isolated from C57BL/10 and C57BL/6 strain mice).



METHODS. As in Fig. 2, with the addition of the following *Va2* oligonucleotides (5' to 3'): 5' *Va2* external, AGCAGCAGGTGAGACAAAGT (amino-acid position,

+1 to +7); 5' *Va2* internal, CCCCAATCTCTGACAGTCTG (amino-acid position, +8 to +14).

in any of the PPD-reactive cells. But the *Vγ2* gene segment has only been found to rearrange the *JCγ2* gene cluster, which differs considerably from the *JCγ4* gene cluster to which the *Vγ1* gene segment generally rearranges. Thus the use of *Vγ1JCγ4* can be dictated as much by the need for the *JCγ4* gene segments as by the need for the *Vγ1* gene segment. It is clear that not all *Vδ* gene segments are permissive for PPD-reactivity, but Table 1 does include at least four that are: *Vδ6.3*, *Vδ6* (p12), *Vδ8* and at least one more unidentified *Vδ*.

The *Vδ6*-expressing hybridomas that we analysed represent all of the *Vδ6*-expressing surface receptor-bearing cells that we have so far generated; they are all derived from newborn thymus and are not a subset selected on the basis of reactivity. The apparent 'predisposition' of these cells to express a specific *γ*-chain gene (26 out of 28 express *Vγ1JCγ4*, see Table 1) indicates that this is another instance of restricted *γ*-*δ*-chain pairing^{12,15,16}. In support of this observation, a *Vδ6*-bearing hybridoma generated by McConnell *et al.*¹² expressed a *Vγ1JCγ4* protein, and several T-cell lines producing *Vδ6* mRNA also expressed the *Vγ1JCγ4* gene product. The tendency for these two chains to be associated with one another could simply be due to physical constraints, or to the ontogenetic timing of *V* gene rearrangements. A third possibility, supported by the fact that the frequently appearing *Vδ6*-*Vγ1JCγ4* combination produces a defined antigen specificity, is that the limited pairing seen is due to a selective pressure to generate this antigen specificity (that is, 'functional constraints').

Finally, the PPD-reactivity of most hybridomas expressing *Vγ1JCγ4* is reminiscent of the ability of certain *Vβ* gene segments to confer reactivity to *Mls* gene products or staphylococ-

cal enterotoxins (the proposed 'superantigens') in 80-90% of the *αβ*⁺T-cells in which they are expressed¹⁷⁻²³. Such *αβ*⁺T-cell responses are unusual in that they seem to involve recognition of relatively nonpolymorphic determinants of MHC gene products. Some *γδ*⁺ T-cell responses involve specific MHC class I molecules (refs 24, 25), or *TL* (ref. 24) or *Qa* (ref. 26) gene products, but the presenting molecule for mycobacterial antigens has so far not been identified. This molecule seems to be relatively nonpolymorphic in some instances^{1,3} but not in others^{4,5}. It remains to be seen whether the use of only one *Vγ* gene in all of the PPD-reactive hybridomas indicates that PPD represents a superantigen for these cells, or that the cells all recognize one particular antigenic epitope or presenting molecule, but the observations presented here should be useful in elucidating the nature of antigen recognition by the *γδ*TCR. □

Note added in proof. A small portion of a probable *Vδ*, isolated from a thymocyte circular DNA library³³, is closely related to the *Vδ8* gene we describe here.

Received 2 August; accepted 19 October 1989.

- O'Brien, R. L. *et al. Cell* **57**, 667-674 (1989).
- Janis, E. M., Kaufman, S. H. E., Schwartz, R. H. & Pardoll, D. M. *Science* **244**, 713-716 (1989).
- Holoshitz, J., König, F., Coligan, J. E., DeBruyn, J. & Strober, S. *Nature* **339**, 226-229 (1989).
- Modlin, R. L. *et al. Nature* **339**, 544-548 (1989).
- Haregewoin, A., Soman, G., Horn, R. C. & Finberg, R. W. *Nature* **340**, 309-312 (1989).
- Marusic-Galesic, S. *et al. J. Immunol.* **142**, 28-34 (1989).
- Iwamoto, A. *et al. J. exp. Med.* **163**, 1203-1212 (1986).
- Chien, Y. *et al. Nature* **330**, 722-727 (1987).
- Yague, J. *et al. Nucleic Acids Res.* **16**, 11355-11364 (1988).
- Arden, B., Klotz, J. L., Siu, G. & Hood, L. E. *Nature* **316**, 783-787 (1985).
- Elliott, J. F., Rock, E. P., Patten, P. A., Davis, M. M. & Chien, Y. *Nature* **331**, 627-631 (1988).
- McConnell, T. J. *et al. J. Immunol.* **142**, 2924-2931 (1989).
- Chien, Y., Iwashima, M., Kaplan, K., Elliott, J. & Davis, M. *Nature* **327**, 677-682 (1987).
- Takagaki, Y., Nakanishi, N., Ishida, I., Kanagawa, O. & Tonegawa, S. *J. Immunol.* **142**, 2112-2121 (1989).
- Korman, A. J., Marusic-Galesic, S., Spencer, D., Kruijsbeek, A. M. & Raulet, D. H. *J. exp. Med.* **168**, 1021-1040 (1988).
- Asanow, D. M. *et al. Cell* **55**, 837-847 (1988).
- Kappler, J. W., Staerz, U., White, J. & Marrack, P. C. *Nature* **332**, 35-40 (1988).
- MacDonald, H. R. *et al. Nature* **332**, 40-45 (1988).
- Kanagawa, O., Palmer, E. & Bill, J. *Cell. Immunol.* **119**, 412-426 (1989).
- Pullen, A., Marrack, P. & Kappler, J. W. *Nature* **335**, 796-801 (1988).
- Abe, R., Vacchio, M. S., Fox, B. & Hodes, R. J. *Nature* **335**, 827-830 (1988).
- White, J. *et al. Cell* **56**, 27-35 (1989).
- Janeway, C. A., Chalupny, J., Conrad, P. J. & Buxner, S. *J. Immunogenet.* **15**, 161-168 (1988).
- Bluestone, J. A., Cron, R. Q., Cotterman, M., Houlden, B. A. & Matis, L. E. *J. exp. Med.* **168**, 1899-1916 (1988).
- Rivas, A., Koide, J., Cleary, M. L. & Engelman, E. G. *J. Immunol.* **142**, 1840-1846 (1989).
- Vidovic, D., Roglic, M., McKune, K., Guerdner, S., MacKay, C. & Dembic, Z. *Nature* **340**, 646-649 (1989).
- Happ, M. P. & Palmer, E. *Eur. J. Immunol.* **19**, 1317-1325 (1989).
- Kubo, R. T. & Pelanne, M. L. *Molec. Immunol.* **20**, 67-76 (1983).
- Oettgen, H. C., Pettley, C. L., Maloy, L. E. & Terhorst, C. *Nature* **320**, 272-275 (1986).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Kumagai, N., Benedict, S., Mills, G. & Gelfand, E. *J. Immunol.* **139**, 1393-1399 (1987).
- Bill, J. *et al. J. exp. Med.* **169**, 115-133 (1989).
- Okazaki, K. & Sakano, H. *EMBO J.* **7**, 1669-1674 (1988).

ACKNOWLEDGEMENTS. We thank Drs Philippa Marrack, William F. Wade and Jerome Bill for reviewing the manuscript and Judy Franconi and Marjorie Greiner for secretarial assistance. This work was supported by funds from the NIH and the American Cancer Society. E.P. is a recipient of a Faculty Research Award from the American Cancer Society. R.L.O.B. is a Howard Hughes Medical Institute associate.

TABLE 1 Summary of *Vγ* and *Vδ* gene usage in PPD-reactive and non-reactive *γδ*⁺ T-cell hybridomas

<i>Vγ</i> → <i>Vδ</i> ↓	PPD-reactive		Non-reactive	
	<i>Vγ1Jγ4</i>	Other <i>Vγ</i>	<i>Vγ1Jγ4</i>	Other <i>Vγ</i>
<i>Vδ6.3</i>	21	0	1	1
<i>Vδ6</i> (p12)	4	0	0	1
<i>Vδ8</i>	1	0	0	2
<i>Vδ1</i>	0	0	0	10
Other <i>Vδ</i>	2	0	2	8
Total	28	0	3	22

Vγ gene usage is classified as *Vγ1Jγ4* or non-*Vγ1Jγ4* as determined by sequencing of PCR amplification products (see Fig. 2 legend). The use of *Vδ6.3*, *Vδ6* (p12) or *Vδ8* was determined first by RNA hybridization analysis² and then by sequencing of PCR amplification products (see Fig. 2 legend). The use of *Vδ1* was determined by RNA hybridization analysis only¹, therefore not all may be rearranged in frame. Unidentified *Vδ* genes are grouped as 'other *Vδ*'.

The colony stimulating factor-1 receptor associates with and activates phosphatidylinositol-3 kinase

Lyuba Varticovski*, Brian Druker†, Deborah Morrison†‡, Lewis Cantley* & Thomas Roberts†

* Department of Physiology, Tufts University School of Medicine, Boston, Massachusetts 02111, USA

† Dana-Farber Cancer Institute, Boston, Massachusetts 02135, USA

COLONY stimulating factor-1 (CSF-1) is a lineage-specific growth factor required for proliferation and survival of mononuclear phagocytes and their precursors. The CSF-1 receptor belongs to a family of ligand-activated protein-tyrosine kinases. Activation of the platelet-derived growth factor receptor, but not the CSF-1 receptor, leads to an increase in phospholipase C activity and a subsequent elevation in intracellular calcium. Recent studies have shown that a novel phosphoinositide (PtdIns) kinase, termed PtdIns-3 kinase, is stimulated by the platelet-derived growth factor receptor and certain oncogenes in the protein-tyrosine kinase family. PtdIns-3 kinase phosphorylates the D-3 hydroxyl position of the inositol ring of PtdIns, and its products do not participate in the generation of the second messenger inositol 1,4,5-trisphosphate (Ins(1,4,5)P₃). Here we report that addition of CSF-1 is followed by activation of PtdIns-3 kinase in a macrophage cell line (P388 D1), which contains CSF-1 receptors, and in BALB/c fibroblasts made to express the human CSF-1 receptor. Furthermore, we show that activation of the CSF-1 receptor results in the accumulation in intact cells of polyphosphoinositides phosphorylated at the D-3 position of the inositol ring. Thus activation of the CSF-1 receptor stimulates PtdIns-3 kinase activity, indicating a novel pathway for CSF-1 receptor-mediated signal transduction.

CSF-1 activity is mediated through a single class of high-affinity receptors¹⁻³. The gene encoding the CSF-1 receptor (*c-fms*) is the normal cellular homologue of the *v-fms* oncogene and belongs to a family of structurally related transmembrane growth-factor receptors with ligand-activated protein-tyrosine kinase activity. Other members of this family include the receptor for platelet-derived growth factor (PDGF) and the *c-kit* gene product, the ligand for which is unknown. These receptors are unique among protein-tyrosine kinases in that their kinase domains contain an insert of 70-100 amino acids. Although the PDGF and CSF-1 receptors are structurally similar, there are differences in the cellular responses that follow their activation. For example, PDGF activates PtdIns turnover and increases intracellular calcium^{4,5}, whereas CSF-1 does not affect PtdIns turnover or cytosolic calcium in murine macrophages⁶.

Recent studies have identified a unique PtdIns kinase activity associated with certain activated tyrosine kinases^{7,8}. This PtdIns kinase, called PtdIns-3 kinase, physically associates with pp60^{v-src} and polyoma middle T/pp60^{c-src} complexes⁹, and is also detected in antiphosphotyrosine immunoprecipitates from cells transformed with *v-src*, *v-fms* and *v-abl*, and in quiescent fibroblasts and smooth muscle cells after stimulation with PDGF^{7,10,11}. PtdIns-3 kinase phosphorylates PtdIns at the D-3 position of the inositol ring of PtdIns to form a novel product, phosphatidylinositol 3-phosphate (PtdIns(3)P)¹².

We investigated whether PtdIns-3 kinase is activated in response to CSF-1. Human CSF-1 receptor was expressed in murine fibroblasts at ~0.5-1 × 10⁴ receptors per cell. Full-length complementary DNA clones for the human CSF-1 receptor gene

and feline *v-fms* were introduced into recombinant retrovirus pLJ (ref. 13). BALB/c fibroblast cell lines expressing these constructs were generated by retroviral infection as previously described¹⁴. Immunoblot analysis was used to screen cells for *fms*-encoded proteins (Fms) and the expression of CSF-1 receptor protein was confirmed by [³⁵S]methionine labelling and immunoprecipitation with monoclonal anti-Fms antibodies. Fibroblasts make CSF-1, but no autocrine effect was observed because murine CSF-1 does not stimulate human CSF-1 receptor, although human CSF-1 stimulates human and murine receptors¹⁵⁻¹⁷. Therefore, no alterations of cell growth or serum requirements were detected in the *c-fms*-expressing cells.

PtdIns kinase activity was detected in immunoprecipitates of CSF-1-stimulated cells (Fig. 1). Confluent serum-starved BALB/c fibroblasts which express CSF-1 receptor were stimulated with human recombinant CSF-1, lysed and immunoprecipitated with anti-phosphotyrosine antibodies. The greatest production of phosphatidylinositol phosphate (PtdInsP) was observed when *c-fms*-expressing cells were stimulated with 20 ng ml⁻¹ of human recombinant CSF-1 (results not

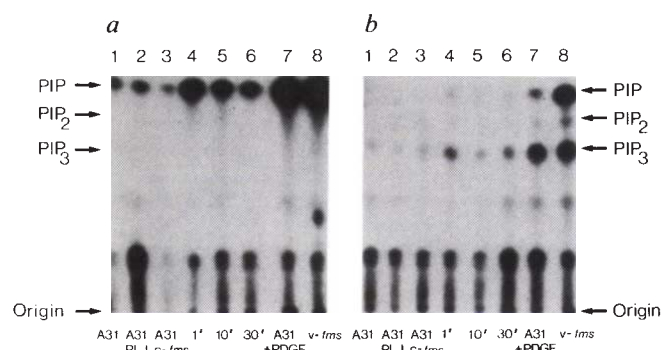


FIG. 1 Time course of CSF-1 dependent activation of PtdIns kinase in anti-phosphotyrosine immunoprecipitates. **a**, TLC analysis of products of lipid kinase assay of anti-phosphotyrosine immunoprecipitates using PtdIns as substrate: A31 parent BALB/c fibroblast cell line (lane 1); A31 cells transfected with the vector alone (lane 2); quiescent A31 fibroblasts expressing human CSF-1 receptor (lane 3); A31 cells expressing human CSF-1 receptor stimulated for 1, 10 and 30 min with 20 ng ml⁻¹ of human recombinant CSF-1 (lanes 4, 5 and 6); A31 cells expressing human CSF-1 receptor stimulated for 10 min with 10 ng ml⁻¹ of recombinant PDGF-BB (lane 7); and *v-fms*-transformed fibroblasts (lane 8). Identification of the product was performed using commercial PtdInsP (PIP) as standard. The migration of the standards (visualized by iodine staining) are indicated by the arrows. Phosphoinositide formation in lanes 1-3 was identified as PtdIns(4)P after deacylation and HPLC analysis. **b**, TLC analysis of the products of lipid kinase assay of anti-phosphotyrosine immunoprecipitates using PtdIns(4,5)P₂ as substrate. Lanes are the same as in **a**. PtdIns and PtdIns(4)P contaminants in the commercial PtdIns(4,5)P₂ resulted in some formation of PIP and PtdInsP₂ (PIP₂). The migration of PtdInsP₃ (PIP₃) standard synthesized from polyoma middle T/pp60^{c-src} immunoprecipitates is indicated by the arrow. **METHODS.** Confluent cultures of BALB/c fibroblasts were placed in serum-free DME medium supplemented with 1% BSA for 24 h to induce quiescence. Recombinant human CSF-1 (courtesy of S. Clark, Genetics Institute) was added to cells for indicated times (prime symbols denote min). Cells were lysed in a 25 mM HEPES buffer, pH 7.5, containing 137 mM NaCl, 1 mM EDTA, 200 μM sodium vanadate, 10% glycerol, 1% Nonidet P-40, 10 μg ml⁻¹ aprotinin and 1 mM phenylmethylsulphonyl fluoride, and immunoprecipitated by incubation for 4 h with anti-phosphotyrosine monoclonal antibody linked to solid support as described by Frackelton *et al.*²⁴. The immunoprecipitates were washed twice with each of the following: (1) ice-cold PBS; (2) 0.5 M LiCl and 0.1 M Tris buffer, pH 7.4; and (3) 0.1 M NaCl, 1 mM EDTA in 0.1 M Tris buffer, pH 7.4 (ref. 9). PtdIns kinase activity was assayed in the immunoprecipitates in the presence of 25 mM HEPES buffer, pH 7.4, and 10 mM MgCl₂. A mixture of 0.1 mg ml⁻¹ PtdIns (**a**) or PIP₂ (**b**) and 0.1 mg ml⁻¹ phosphatidylserine, dispersed by zonation in 25 mM HEPES buffer pH 7.4 and 1 mM EDTA, was added to the immunoprecipitates. The reaction was initiated by addition of 20 μCi of [³²P]ATP and 100 μM ATP as previously described¹². The products were separated by TLC in 1-propanol:acetic acid (2N) [65:35 (v/v)], visualized by autoradiography and identified by comparison with non-labelled standards.

‡ Present address: Department of Medicine, Howard Hughes Medical Institute, University California, San Francisco, California 94143, USA.

shown). This concentration of CSF-1 was used in all subsequent experiments, and is equivalent to $3,000 \text{ U ml}^{-1}$ (1U is equivalent to 0.44 fmol); and is comparable to the amount of CSF-1 that elicits a maximal mitogenic response in murine macrophages which express a similar number of CSF-1 receptors^{18,19}. Increased immunoprecipitable PtdIns kinase activity was detected

within 1 min after addition of CSF-1 (Fig. 1a). No CSF-1-dependent PtdIns-3 kinase activity was detected when parent BALB/c fibroblasts or cells infected with the vector alone were used (results not shown). The PtdIns-3 kinase activity detected in antiphosphotyrosine immunoprecipitates of PDGF-stimulated cells was fivefold higher than that detected in CSF-1-stimulated cells (lanes 5 and 7, Fig. 1a). These activities correspond to the relative abundance of each receptor. Antiphosphotyrosine immunoprecipitates of *v-fms*-transformed cells had elevated PtdIns-3 kinase activity in the absence of exogenous CSF-1 (lane 8, Fig. 1a).

Antiphosphotyrosine-antibody immunoprecipitates from CSF-1-stimulated cells also had lipid kinase activities which used phosphatidylinositol 4-phosphate (PtdIns(4)P) and phosphatidylinositol 4,5-bisphosphate (PtdIns(4,5)P₂) as substrates. Figure 1b presents the results of an experiment in which PtdIns(4,5)P₂ was used as a substrate. The product of this reaction migrated with phosphatidylinositol trisphosphate (PtdInsP₃) standard on TLC. A significant increase in PtdIns(4,5)P₂ kinase activity was detectable within 1 min (the decrease in activity seen at 10 min in Fig. 1b, lane 5, was not reproduced in three other experiments). Elevated PtdIns(4,5)P₂ kinase activity was also detected in immunoprecipitates from *v-fms*-transformed cells and in BALB/c control cells stimulated with PDGF. When PtdIns(4)P was used as a substrate, a PtdInsP₂ was produced with a similar CSF-1 dependence (data not shown). HPLC identification of the products obtained in these reactions is shown in Fig. 2. The lipids were deacylated and identified by comparison of their HPLC-migration positions with those of standards: [³²P]PtdIns(3)P, [³²P]PtdInsP₂ and [³²P]PtdInsP₃ standards were synthesized using middle T/pp60^{c-src} immunoprecipitates¹⁹; [³H]PtdIns(4)P and [³H]PtdIns(4,5)P₂ standards were from a commercial source. No D-3 phosphorylated polyphosphoinositides were detected in control lanes. The deacylation product of the PtdInsP which was produced by antiphosphotyrosine immunoprecipitates from CSF-1-stimulated cells comigrated with glycerophosphoinositol 3-phosphate (GroPIns(3)P); the deacylated PtdInsP₂ comigrated with glycerophosphoinositol 3,4-bisphosphate (GroPIns(3,4)P₂) standard, and the deacylated PtdInsP₃ comigrated with glycerophosphoinositol-trisphosphate (GroPInsP₃) standard. Thus, CSF-1

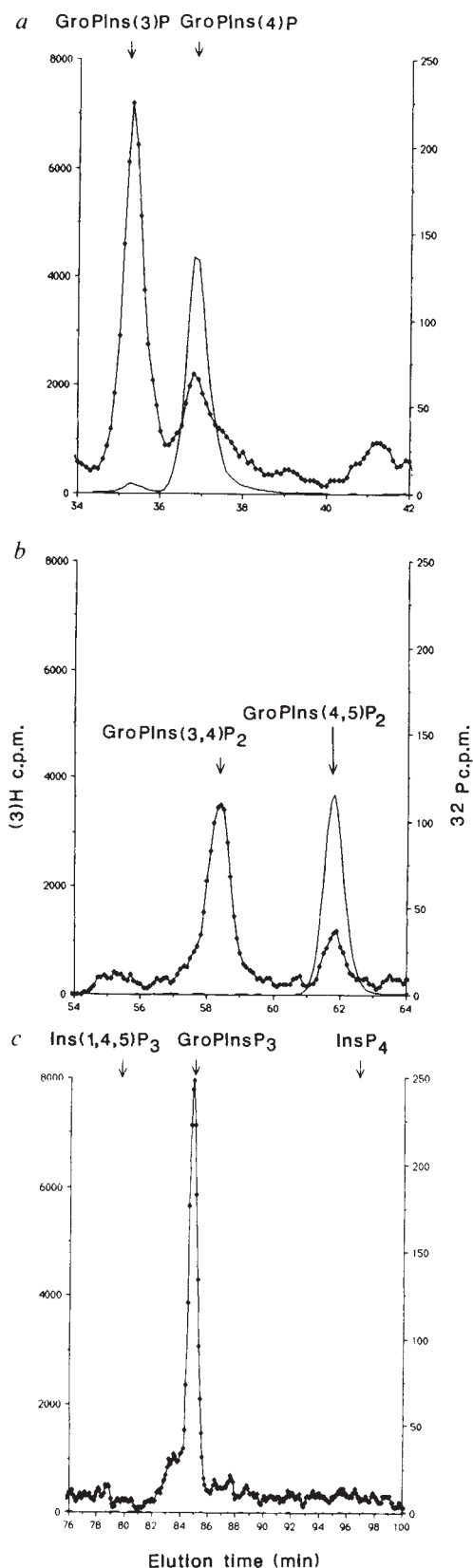


FIG. 2 HPLC identification of polyphosphoinositides generated from antiphosphotyrosine immunoprecipitates of CSF-1-stimulated cells. a, HPLC separation of the deacylated product of PtdIns kinase assay of anti-phosphotyrosine immunoprecipitates from CSF-1-stimulated cell. The [³²P]PtdInsP was produced and deacylated with methylamine reagent as previously described¹⁹. Solid line represents [³H]GroPIns(4)P standard. ■, Deacylated ³²P-labelled PtdInsP generated from fibroblasts which express human CSF-1 receptor after exposure to 20 ng ml^{-1} of human recombinant CSF-1 for 5 min. The migration position of GroPIns(3)P was determined from a parallel run of [³²P]GroPIns(3)P standard generated from middle T/pp60^{c-src} immunoprecipitates. b, HPLC separation of the deacylated product generated as above except using PtdIns(4)P as substrate. Solid line represents [³H]GroPIns(4,5)P₂ standard. Closed squares represent deacylated ³²P-labelled PtdInsP₂. The migration position of the GroPIns(3,4)P₂ was determined from a parallel run of [³²P]GroPIns(3,4)P₂ standard generated from middle T/pp60^{c-src} immunoprecipitates. c, HPLC separation of the deacylated product generated as above, except using PtdIns(4,5)P₂ as substrate. Closed squares represent deacylated ³²P-labelled PtdInsP₃. Arrows indicate the migration position from a parallel run of [³H]Ins1,4,5P₃ and [³H]Ins1,3,4,5P₄ standards (NEN), and the position of [³²P]GroPInsP₃ generated from middle T/pp60^{c-src} immunoprecipitates.

METHODS. Deacylated phospholipids were suspended in $400 \mu\text{l H}_2\text{O}$ with appropriate standards and separated by anion exchange HPLC using a Partisphere Sax column with a gradient from 0 to $1.0 \text{ M (NH}_4\text{)}_2\text{HPO}_4$ (pH 3.8) developed over 130 min^{11} . Isotope detection was performed using an on-line continuous flow scintillation detector (Radiomatic Instruments) which separates ³H, ³²P counts, and a parallel reading of UV channel for detection of non-labelled ADP and ATP standards.

receptor seems to stimulate the same lipid kinase activities as those associated with middle T/pp60^{c-src} and PDGF receptor¹⁹.

A monoclonal antibody which specifically recognizes the N-terminal domain of both *c-fms* and *v-fms* proteins was used to investigate whether PtdIns-3 kinase directly associates with the CSF-1 receptor. Increased phosphorylation of PtdIns was detected in immunoprecipitates from CSF-1-stimulated cells when anti-Fms antibodies were used. HPLC analysis of the deacylated lipid showed that although some PtdIns(4)P was produced by the immunoprecipitates, the principal lipid was PtdIns(3)P. A similar result was detected using immunoprecipitates from *v-fms*-transformed cells (Fig. 3). These results indicate that PtdIns-3 kinase is physically associated with ligand-activated CSF-1 receptor and with *v-fms* oncoprotein.

We also investigated the ability of CSF-1 to stimulate PtdIns-3 kinase in a cell line which expresses endogenous CSF-1 receptor. P388 D1 is a murine macrophage cell line which expresses 3.5×10^3 CSF-1 receptors per cell². After addition of recombinant human CSF-1, an increase in PtdIns-3 kinase activity was detected in anti-phosphotyrosine immunoprecipitates (data not shown). Thus, activation observed with the transfected receptor is not merely a consequence of expression of the receptor in fibroblasts.

The polyphosphoinositides produced by PtdIns-3 kinase in immunoprecipitates from CSF-1-stimulated cells were also elevated in response to CSF-1 in intact cells. P388 D1 macrophage cells were starved of serum for 14 h and labelled with $^{32}\text{PO}_4^{3-}$ for 3 h. Analysis of deacylated phospholipids from starved murine macrophage cells showed small but detectable amounts of ^{32}P -labelled lipids which comigrated with deacylated PtdIns(3)P, PtdIns(3,4)P₂ and PtdInsP₃ standards. These lipids each represented, respectively, 3.2%, 0.7% and 0.2% of the ^{32}P label incorporated into PtdIns(4)P in the same HPLC analysis.

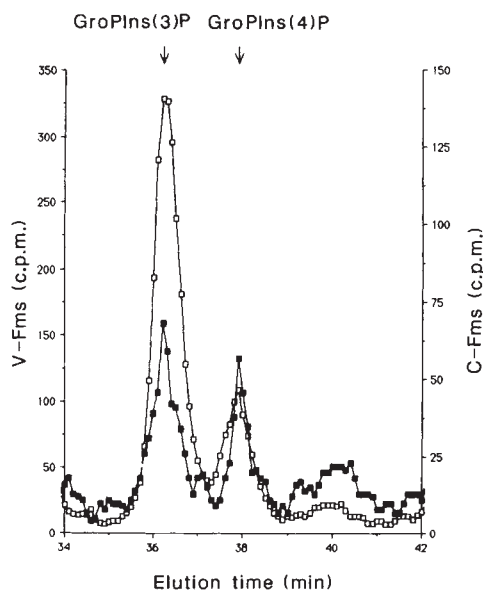


FIG. 3 HPLC analysis of the deacylated PtdInsP product produced by the anti-Fms immunoprecipitates of *c-fms* expressing fibroblasts which were stimulated with CSF-1- and *v-fms*-transformed cells. ■, Deacylated ^{32}P -labelled PtdInsP generated from immunoprecipitates of *c-fms*-expressing cells stimulated with CSF-1. □, Deacylated ^{32}P -labelled PtdInsP generated from immunoprecipitates of *v-fms*-transformed cells (no CSF-1 added). CSF-1-stimulated fibroblasts which express the CSF-1 receptor and *v-fms*-transformed cells were lysed, equalized for the amount of protein, and reacted with rat monoclonal anti-Fms antibody for 6 h. An anti-rat polyclonal antibody linked to a solid support (Cooper Diagnostics) was used as a second antibody. PtdInsP generated by immunoprecipitable PtdIns kinase was analysed as indicated in Fig. 2.

This observation differs from results obtained from serum-starved fibroblasts and smooth muscle cells where no PtdIns(3,4)P₂ or PtdInsP₃ was detected¹⁹, and could be due to the difficulty in synchronizing these macrophages into a G₀ state. Addition of human recombinant CSF-1 for 5 min resulted in a 70% increase in PtdIns(3,4)P₂, and a 300% increase in PtdInsP₃. The levels of these lipids declined but remained above control values at 30 min (Fig. 4). No CSF-1-mediated changes in the content of PtdIns(3)P were seen. Similar results were observed in smooth muscle cells¹¹ after addition of PDGF, suggesting that intracellular PtdIns(3)P levels do not change during mitogenic response.

The physiological significance of PtdIns-3 kinase remains to be determined. None of the products of this kinase are in the pathway for generating the Ca^{2+} -regulating second messenger, Ins(1,4,5)P₃, suggesting that these lipids provide signals independent of the conventional PtdIns turnover response. Evidence for the importance of PtdIns-3 kinase for growth and transformation is provided by studies of mutants for the polyoma middle T gene and mutants for the PDGF receptor gene. In most middle T gene mutants which are defective in cell transformation, PtdIns-3 kinase fails to associate with the middle T/pp60^{c-src} complex^{9,20}. Included in this group of transformation-defective genes are mutants in which the middle T protein associates with and activates the protein-tyrosine kinase activity of pp60^{c-src}, but fails to associate with or activate PtdIns-3 kinase. A similar phenotype has been observed for a mutant PDGF receptor in which 83 of the 100 amino acids in the protein-

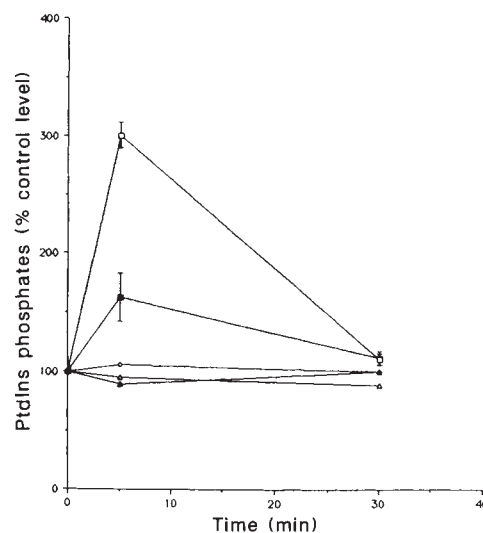


FIG. 4 Time course of the levels of polyphosphoinositides in intact murine macrophages after stimulation with CSF-1: *In vivo* levels of PtdInsP₃ (□); PtdIns(3,4)P₂ (■); PtdIns(3)P (△); PtdIns(4)P (◇) and PtdIns(4,5)P₂ (▲) are expressed as percentages of the control level. Each point represents three independent experiments. The error bars represent s.d. Murine macrophages were grown to confluency in RPMI, 10% FCS and deprived of serum for 14 h. Cells were washed with phosphate-free MEM supplemented with 1-glutamine and 1% BSA, and labelled for 3 h with 0.1 mCi $^{32}\text{PO}_4^{3-}$ ml⁻¹. The medium was removed, cells were rinsed with the same medium without $^{32}\text{PO}_4^{3-}$, and stimulated with 10 ng ml⁻¹ human recombinant CSF-1 at time = 0. At the end of the periods indicated, the medium was removed and 1 M HCl was added to the plate. Cells were gathered with the aid of a Teflon cell scraper and the phospholipids were extracted into methanol:chloroform (1:2, v/v)¹¹. The phospholipids were deacylated as indicated in the legend to Fig. 2 and identified on HPLC by comparison with deacylated standards. Between 0.8 and 2×10^6 ^{32}P c.p.m. were used per HPLC run. In a typical experiment (1×10^6 c.p.m. injected into HPLC), the total counts incorporated at time zero were 177,206 for PtdIns(4)P, 219,936 for PtdIns(4,5)P₂, 5,818 for PtdIns(3)P, 1,252 for PtdIns(3,4)P₂, and 366 for PtdInsP₃.

tyrosine kinase domain insert are deleted^{21,22}. This receptor retained PDGF-stimulated protein-tyrosine kinase activity, stimulated PtdIns turnover and calcium elevation, but did not associate with PtdIns-3 kinase and failed to mediate PDGF-dependent DNA synthesis. These results indicate that activation of PtdIns-3 kinase could be essential for the growth and transformation response of these genes. But transformation-defective mutants of middle T protein (ref. 23, and B.D. and T.R., unpublished observations) and of p60^{v-src} (ref. 20) which retain wild-type levels of PtdIns-3 kinase have been found, suggesting that this activation alone is not sufficient for transformation. We have shown that the CSF-1 receptor has similar properties to the PDGF receptor with respect to activation of the PtdIns-3 kinase pathway, indicating that production of D-3 phosphorylated phosphoinositides could be critical for the growth-stimulating activities of these two receptors. □

Received 24 July; accepted 12 October 1989.

1. Stanley, E. R. & Jubinsky, P. T. *Clin. Haemat.* **13**, 329–348 (1984).
2. Guilbert, L. J. & Stanley, E. R. *J. Biol. Chem.* **261**, 4024–4032 (1986).
3. Guilbert, L. J., Tynan, P. W. & Stanley, E. R. *J. cell. Biochem.* **31**, 203–216 (1986).
4. Habenicht, A. J. R. et al. *J. Biol. Chem.* **256**, 12329–12335 (1981).
5. Berridge, M., Heslop, J., Irvine, R. F. & Brown, K. D. *Biochem. J.* **222**, 195–201 (1984).
6. Whetton, A. D., Monk, P. N., Consalvey, S. D. & Downes, C. P. *EMBO J.* **5**, 3281–3286 (1986).
7. Whitman, M., Kaplan, D., Roberts, T. & Cantley, L. *Biochem. J.* **247**, 165–174 (1987).
8. Kaplan, D. R. et al. *Cell* **50**, 1021–1029 (1987).
9. Whitman, M., Kaplan, D., R. Schaffhausen, B., Cantley, L. and Roberts, T. *Nature* **315**, 239–242 (1985).
10. Varticovski, L. & Cantley, L. *Blood* **72**, 232a (1988).
11. Auger, K. R., Serunian, L. A., Soltoff, S. P., Libby, P. & Cantley, L. *Cell* **57**, 167–175 (1989).
12. Whitman, M., Downes, C. P., Keeler, M., Keller, T. & Cantley, L. *Nature* **332**, 644–646 (1988).
13. Morrison, D. K., Browning, P., White, M. & Roberts, T. *Molec. cell. Biol.* **8**, 176–185 (1988).

14. Pivnicka-Worms, H., Saunders, K. B., Roberts, T. M., Smith, A. E. & Cheng, S. H. *Cell* **49**, 75–82 (1987).
15. Das, S. K., Stanley, E. R., Guilbert, L. J. & Forman, L. W. *J. cell. Physiol.* **104**, 359–364 (1980).
16. Russel, M. F. et al. *Nature* **325**, 549–552 (1987).
17. Rousel, M. F., Downing, J. R., Rettenmier, C. W. & Sherr, C. J. *Cell* **55**, 979–986 (1988).
18. Stanley, E. R. & Guilbert, L. J. *J. Immunol. Meth.* **42**, 253–284 (1981).
19. Morgan, C., Pollard, J. W. & Stanley, R. J. *cell. Physiol.* **130**, 420–427 (1987).
20. Fukui, Y. & Hanafusa, H. *Molec. cell. Biol.* **9**, 1651–1658 (1989).
21. Coughlin, S. R., Escobedo, J. A. & Williams, L. T. *Science* **243**, 1191–1193 (1989).
22. Williams, L. T. *Science* **243**, 1564–1570 (1989).
23. Morgan, W. C., Kaplan, D. R., Pallas, D. C. & Roberts, T. M. *J. Virol.* **62**, 3407–3411 (1988).
24. Frackleton, A. R., Jr., Ross, A. H. & Eisen, N. H. *Molec. cell. Biol.* **3**, 1343–1352 (1983).

ACKNOWLEDGEMENTS. We thank I. Arias and H. Pivnicka-Worms for helpful suggestions and M. Keeler for assistance with HPLC analysis. This work was supported by the NIH.

Differential expression of genes encoding α , β and γ retinoic acid receptors and CRABP in the developing limbs of the mouse

Pascal Dollé, Esther Ruberte, Philippe Kastner, Martin Petkovich, Carol M. Stoner*, Lorraine J. Gudas* & Pierre Chambon†

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de médecine, 11 rue Humann, 67085 Strasbourg Cédex, France

* Department of Biological Chemistry and Pharmacology, Harvard Medical School, and Dana-Farber Cancer Institute, Boston, Massachusetts 02115, USA

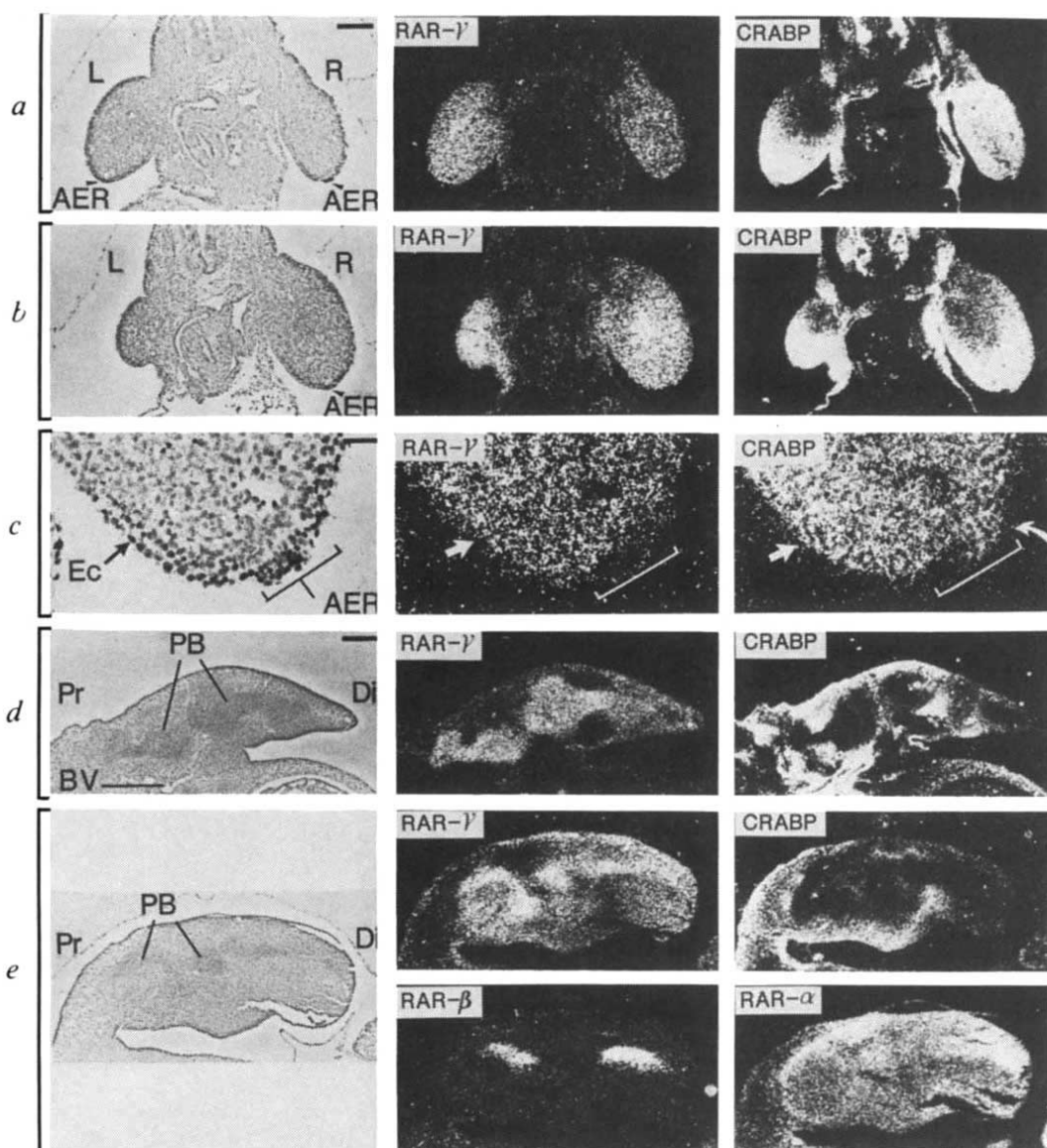
RETINOIC acid has profound effects on vertebrate limb morphogenesis (refs 1–6, reviewed in refs 7–9), including in the mouse, where it can act as a teratogen generating phocomelia and bone defects^{10–12}. A retinoic acid gradient^{13–15}, possibly amplified by a graded distribution of cellular retinoic acid-binding protein (CRABP)¹⁶, could provide positional information across the antero-posterior axis of the chick limb bud. The discovery of nuclear retinoic acid receptors (RARs)^{17–22} acting as retinoic acid-inducible enhancer factors (reviewed in refs 23, 24) provided a basis for understanding how retinoic acid signals could be transduced at the level of gene expression²⁵. We have now used *in situ* hybridization to study the distribution of messenger RNA transcripts of the three murine receptors (mRARs) and CRABP during mouse limb development. Both mRAR α and mRAR γ transcripts, but not those for mRAR β , are present and uniformly distributed in the limb bud at day 10 post-coitum, whereas CRABP transcripts have a graded proximo-distal distribution, indicating that differential expression of CRABP, but not of mRAR α or mRAR γ , could participate in the establishment of the morphogenetic field. At later stages, mRAR γ transcripts become specific to the cartilage cell lineage and to the differentiating skin and mRAR β transcripts are mostly restricted to the interdigital mesenchyme. CRABP transcripts, however, are excluded from regions expressing mRAR γ and mRAR β . These results indicate that all three RARs and CRABP have specific functions during morphogenesis and differentiation of the mouse limb.

We synthesized ³⁵S-labelled antisense RNA probes from mRAR α , mRAR β and mRAR γ (ref. 21) and murine CRABP (ref. 26) complementary DNA clones, and hybridized them to serial sets of contiguous histological sections. Three stages of the mouse limb development were studied, corresponding to nondifferentiated limb buds (day 10 post-coitum (p.c.)), the appearance of precartilaginous blastemas (day 12.5 p.c.) and the beginning of ossification and digit separation (day 14.5 p.c.). Forelimb bud sections at day 10 p.c. showed that mRAR γ (Fig. 1a, b) and mRAR α (data not shown) transcripts were homogeneously distributed throughout the limb bud. The signal for mRAR γ transcripts stopped at the proximal limit of the buds, whereas that for mRAR α transcripts extended into adjacent tissues. mRAR β transcripts were detected only in the most proximal region of the limb buds and in the mesenchyme of the flank (data not shown). Whereas CRABP transcripts were abundant and uniformly distributed in posterior sections of the forelimb (Fig. 1a, R), there was a gradual proximal-to-distal increase in transcripts in more anterior sections (Fig. 1a, L; b), so that no labelling was seen in the proximal core of the bud, whereas it was maximal at the distal tip. No CRABP transcripts were detected in the ectodermal layer (Ec), except in the apical ectodermal ridge (AER), where only the most dorsal cells were repeatedly labelled, irrespective of their antero-posterior location (Fig. 1a–c). By contrast, both the AER and ectoderm were devoid of mRAR γ transcripts (Fig. 1c). The immunohistochemistry study of Maden *et al.*¹⁶ has similarly revealed a higher concentration of CRABP in the distal progress zone of the chick limb bud. Moreover, the existence of an antero-posterior gradient of CRABP in this zone, with opposite polarity to that of retinoic acid¹³ has been reported¹⁶. We did not find an equivalent gradient of CRABP transcripts. Further studies are required to establish whether such a gradient of CRABP exists. The results described here, however, indicate that there could be a proximo-distal gradient of CRABP, which is maximal distally. This could create a gradient of free retinoic acid across the limb bud, which may be interpreted by the uniformly expressed mRAR α or mRAR γ , or both, at the time when positional information is to be specified.

By day 12.5 p.c., a remarkable segregation of the domains of expression of mRAR γ and CRABP was apparent. Murine RAR γ transcripts were proximally restricted to central precartilaginous blastemas (PB) of both forelimb (Fig. 1d) and hindlimb (Fig. 1e), and more homogeneously distributed in the distal undifferentiated mesenchyme of both limbs. By marked contrast, CRABP transcripts appeared to be excluded from the central PB and were most abundant in cells underlying the ectoderm

† To whom correspondence should be addressed.

FIG. 1 Detection of mRARs and CRABP transcripts in the developing murine limbs by gestational days 10 and 12.5 p.c. Bright-field views of selected section planes are shown in left panels for histological identification. The right and middle panels are dark-field views showing autoradiography signal (white grains) on contiguous sections hybridized with mRARs and CRABP probes, as indicated. Sections hybridized to CRABP probes were exposed half as long as all others. *a, b*, Two serial transverse sections across the forelimb buds of a 10-day-p.c. embryo. The section shown in *b* is more anterior than in that shown in *a*. Because the section plane is slightly biased, the left limb (L) is cut more anteriorly. Bar, 160 μ m. *c*, Higher magnification of the distal region of the forelimb bud (including the AER as indicated) of a 10-day-p.c. embryo. Note that absence of labelling in the ectoderm (Ec) with both probes (the short white arrows point to the same Ec position in the dark-field views as the black arrow in the bright-field view), and the selective labelling of the dorsal part of the AER with the CRABP probe (curved arrow). Bar, 50 μ m. *d*, Longitudinal section of a 12.5-day-p.c. forelimb. Bar, 200 μ m. *e*, Sagittal section of a 12.5-day-p.c. hindlimb extremity; the footplate points to the right. L, left limb bud; R, right



limb bud; AER, apical ectodermal ridge; Ec, ectoderm; Pr, proximal; Di, distal; PB, precartilaginous blastemas; BV, blood vessels.

METHODS. Murine RAR α , mRAR β , mRAR γ (ref. 21) and CRABP (ref. 26) full-length cDNAs subcloned into plasmid pSG5 (ref. 34) in an antisense orientation were used as templates in a T7 polymerase *in vitro* transcription reaction. α - 35 S-labelled CTP (Amersham) was added to the reaction to yield antisense RNA probes of specific activity $\sim 5.10^8$ c.p.m. μ g $^{-1}$. Probe average length was reduced to ~ 100 nucleotides by limited alkaline hydrolysis. Collection, paraformaldehyde fixation, paraffin embedding and sectioning of mouse embryos were carried out as described in ref. 35. The *in situ*

(Fig. 1*d, e*) and surrounding the limb axial blood vessels (Fig. 1*d, BV*), with very little CRABP expression in the distal mesenchyme where mRAR γ transcripts were homogeneously distributed. At the same stage, mRAR α transcripts were widely distributed (Fig. 1*e*), whereas mRAR β transcripts (Fig. 1*e*) were only expressed in a discrete proximal region of the hindlimb mesenchyme, distinct but adjacent to the zeugopod PB, and in several regions of loosely packed cells in the forefoot and hindfoot plates, similar to that shown in Fig. 1*e*.

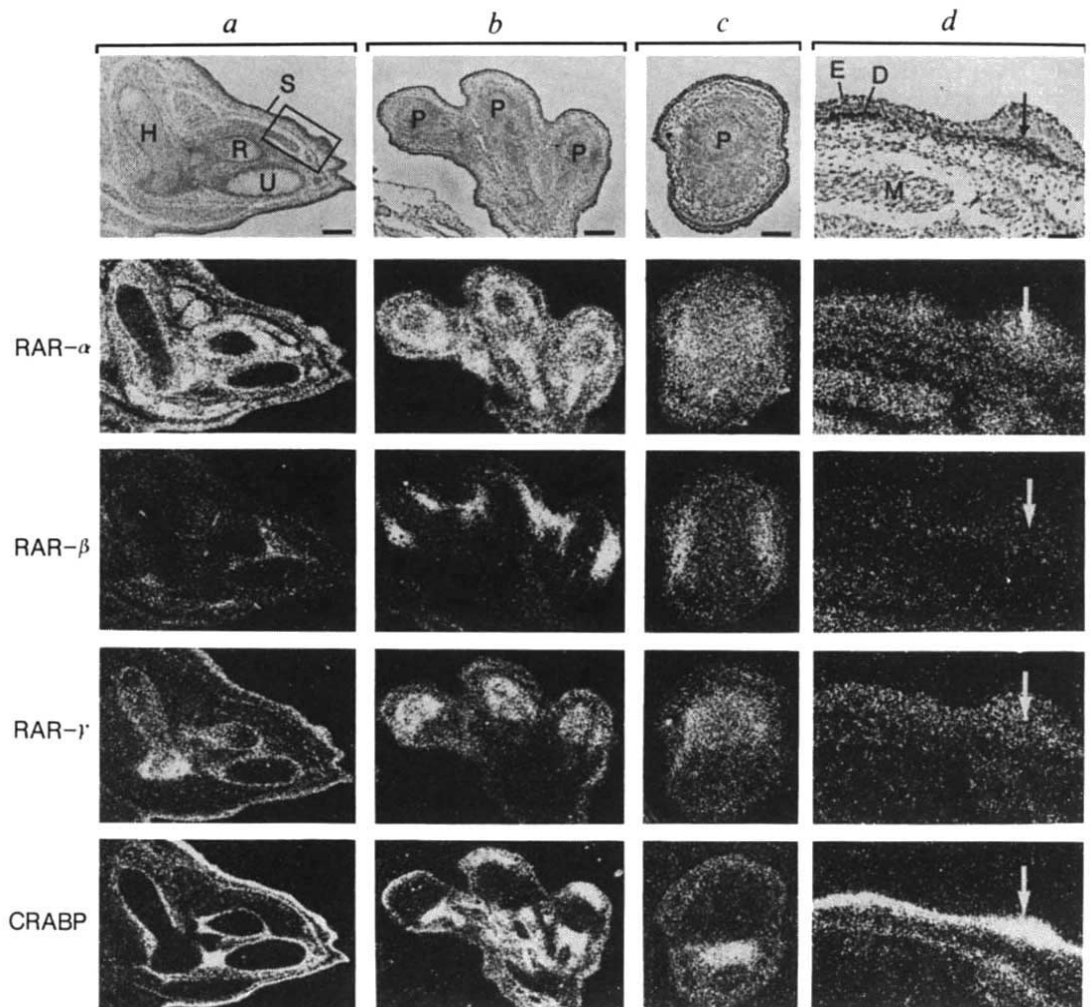
By day 14.5 p.c., mRAR γ transcripts were found exclusively in the epidermal and dermal layers of the fetal skin (Fig. 2*a, d*) and in the cartilage cells of every bone model (Fig. 2*a, b*), with maximal expression in their growing extremities (for example, that of humerus in Fig. 2*a*) and no significant expression in ossification centres. Nevertheless, mRAR γ labelling remained

hybridization protocol was previously described³⁵, except that the prehybridization step was omitted. Contiguous sections (5- μ m thick) were hybridized respectively to each of the four probes, in a sequentially repeated manner so that 50- μ m regularly spaced section planes were analysed. Emulsion autoradiography exposures were for 14 days for mRARs probes, and for 7 days for the CRABP probe. Specificity controls included hybridization of parallel sections to α - 35 S-labelled dATP tailed oligonucleotide probes corresponding to regions of least homology across the mRARs. These probes resulted in the same signal distribution as that obtained with riboprobes, but longer exposure times were required.

more diffuse in the extremities of the fingers (Fig. 2*c*) where bones, muscles and skin had not yet differentiated (see legend to Fig. 2). By contrast, the mRAR α gene was expressed in all tissues including muscle, skin and perichondrium, but at lower levels in cartilage (Fig. 2*a-d*). Murine RAR β transcripts were highly restricted to mesenchymal cells located in the interdigital regions (Fig. 2*b, c*). CRABP transcripts were selectively distributed at the periphery of cartilaginous bone models and in the dermis, but not in the epidermal layer, nor in muscles (Fig. 2*a-d*). It is interesting that CRABP transcripts were not detected in the precise regions of the digits where the genes encoding mRAR β and mRAR γ were maximally expressed, although they were clearly seen in the surrounding mesenchyme (Fig. 2*b, c*).

The fact that the genes encoding the three RARs are differentially expressed at the time of cytodifferentiation in the murine

FIG. 2 Distribution of mRARs and CRABP transcripts in the limbs by gestational day 14.5 p.c. Top panels are bright-field views and lower panels show dark-field sections hybridized to mRAR α , mRAR β , mRAR γ and CRABP probes as indicated. *a*, Sagittal section of the elbow region, crossing the fetal humerus, radius and ulna. Note that at this stage of development the distal region of the humerus (H) and the proximal region of the radius (R) which are seen in this section contain both cartilaginous areas (lower part in the humerus, left part in the radius) and an ossification centre (upper part in the humerus, right part in the radius), whereas the section crosses the ulna essentially in an ossification centre. Bar, 200 μ m. *b*, Section across the extremity of the hindlimb, crossing three digits. Bar, 160 μ m. *c*, Transverse section of a forelimb finger extremity, just distal to the interdigital cleft. The distal regions of the fingers are still poorly differentiated; this is particularly clear in the case of the skin where the epidermis and dermis are not yet differentiated (compare with upper panel of *d*) and for the central condensation where chondrogenesis is not yet complete. The mRAR β signal is restricted to the mesenchyme facing the adjacent digits, whereas the mRAR γ signal is more evenly distributed, although enhanced in the central condensation. Bar, 100 μ m. *d*, Higher magnification of the region of the forelimb boxed in *a*. The arrows point to the basal lamina separating the dermis (D) and the epidermis (E); mRAR α labelling appears ubiquitous (including muscles (M),



whereas no significant mRAR β labelling is seen, mRAR γ appears to be expressed both in dermis and epidermis, and CRABP is expressed in the dermis to such a great extent that the autoradiography grain is visible on the bright-field view (top panel), but not in the epidermis or in muscles. Bar, 75 μ m. H, humerus; R, radius; U, ulna; S, skin; P, phalangeal bone anlagen; E, epidermis; D, dermis; M, muscle.

limb, strongly supports the idea that each RAR performs specific functions^{21,22}. The specific localization of mRAR γ transcripts in the cartilaginous cell lineage and in the differentiating skin, which extends throughout the embryo²⁷, indicates that mRAR γ could mediate the effects of retinoic acid on chondrogenesis^{11,28,29} and on skin differentiation³⁰. The expression of the gene encoding mRAR β could be associated with programmed cell death³¹, because it is highly restricted to the mesenchyme of the interdigital regions at the time of digit separation, and to their possible precursor areas in the differentiating footplates. CRABP transcripts are generally excluded from the areas expressing the mRAR β and mRAR γ genes (although they are present in the surrounding regions) and to a large extent from those areas expressing the gene encoding mRAR α . Thus, it seems possible that CRABP is not directly involved in the transduction of the retinoic acid signal by RARs. Whether it has a role in the control of free retinoic acid levels as previously suggested by its gradient distribution in the chick limb bud¹⁶ is unknown. It is noteworthy that CRABP transcripts are preferentially localized in the skin dermal layer, whereas the genes encoding mRAR α and mRAR γ are expressed in both dermis and epidermis, because skin differentiation could be programmed by progressive migration of keratinocytes away from the dermal vascular source of retinoids^{32,33}. □

Received 21 August; accepted 31 October 1989.

1. Maden, M. *Nature* **295**, 672-675 (1982).
2. Thoms, S. D. & Stocum, D. L. *Dev Biol.* **103**, 319-328 (1984).
3. Scadding, S. R. & Maden, M. *J. embryol. exp. Morphol.* **91**, 35-53 (1986).
4. Tickle, C., Alberts, B. M., Wolpert, L. & Lee, J. *Nature* **296**, 564-565 (1982).
5. Summerbell, D. *J. Embryol. exp. Morphol.* **78**, 269-289 (1983).
6. Tickle, C., Lee, J. & Eichele, G. *Dev Biol.* **109**, 82-95 (1985).
7. Maden, M. *Trends Genet.* **1**, 103-107 (1985).
8. Brookes, J. P. *Neuron* **2**, 1285-1294 (1989).
9. Eichele, G. *Trends Genet.* **5**, 246-251 (1989).
10. Kochhar, D. M. *Teratology* **7**, 289-299 (1973).
11. Satre, M. A. & Kochhar, D. M. *Dev Biol.* **133**, 529-536 (1989).
12. Sulik, K. K., Johnston, M. C. & Dehart, D. B. *Teratology* **35**, 32A (1987).
13. Thaller, C. & Eichele, G. *Nature* **327**, 625-628 (1987).
14. Slack, J. M. W. *Nature* **327**, 553-554 (1987).
15. Thaller, C. & Eichele, G. *Development* **103**, 473-483 (1988).
16. Maden, M., Ong, D. E., Summerbell, D. & Chytil, F. *Nature* **335**, 733-735 (1988).
17. Petkovich, M., Brand, N. J., Krust, A. & Chambon, P. *Nature* **330**, 444-450 (1987).
18. Giguere, V., Ong, E. S., Segui, P. & Evans, R. M. *Nature* **330**, 624-629 (1987).
19. Brand, N. *et al.* *Nature* **332**, 850-853 (1988).
20. Benbrook, D., Lernhardt, E. & Pfahl, M. *Nature* **333**, 669-672 (1988).
21. Zelent, A., Krust, A., Petkovich, M., Kastner, P. & Chambon, P. *Nature* **339**, 714-717 (1989).
22. Krust, A., Kastner, P., Petkovich, M., Zelent, A. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **86**, 5310-5314 (1989).
23. Green, S. & Chambon, P. *Trends genet.* **4**, 309-314 (1988).
24. Evans, R. M. *Science* **240**, 889-895 (1988).
25. Robertson, M. *Nature* **330**, 420-421 (1987).
26. Stoner, C. M. & Gudas, L. J. *Cancer Res.* **49**, 1497-1504 (1989).
27. Ruberte, E. *et al.* *Development* (in the press).
28. Galland, F. & Kistler, A. *Wilhelm Roux Arch. dev. Biol.* **189**, 25-33 (1980).
29. Kochhar, D. M., Penner, J. D. & Tellone, C. *Teratol. Carcinog. Mutagen.* **4**, 377-387 (1984).
30. Shapiro, S. S. In *Retinoids and Cell Differentiation* (ed. H. I. Sherman) 30-55 (1985).

31. Menkes, B. & Ilies, A. *Revue rom. embryol. Cytol.* **2**, 161-172 (1965).
32. Kapan, R., Traska, G. & Fuchs, E. *J. cell. biol.* **105**, 427-440 (1987).
33. Kapan, R. & Fuchs, E. *J. cell. Biol.* **109**, 295-307 (1989).
34. Green, S., Issemann, I. & Scheer, E. *Nucleic Acids Res.* **16**, 369-372 (1988).
35. Dollé, P. & Duboule, D. *Embo J.* **8**, 1507-1515 (1989).

ACKNOWLEDGEMENTS. We thank Drs D. Duboule and G. Morriss-Kay for discussions and critical reading of the manuscript, C. Werlé, A. Landmann and J. M. Lafontaine for illustrations, and the secretarial staff. This work was supported by the INSERM, the CNRS, the Fondation pour la Recherche Médicale and the Association pour la Recherche sur le Cancer. M.P. was supported by a fellowship from the FRMF.

Mutations in the *p53* gene occur in diverse human tumour types

Janice M. Nigro*, Suzanne J. Baker*,
Antonette C. Preisinger*, J. Milburn Jessup†,
Richard Hostetter†, Karen Cleary†, Sandra H. Bigner‡,
Nancy Davidson*, Stephen Baylin*, Peter Devilee§,
Thomas Glover||, Francis S. Collins||, Ainsley Weston¶,
Rama Modali¶, Curtis C. Harris¶ & Bert Vogelstein*#

* The Johns Hopkins Oncology Center, 424 North Bond Street, Baltimore, Maryland 21231, USA

† M. D. Anderson Hospital, Houston, Texas 77030, USA

‡ Department of Pathology, Duke University, Durham, North Carolina 27710, USA

§ Department of Human Genetics, University of Leiden, Holland

|| Department of Human Genetics, University of Michigan, Ann Arbor, Michigan 48109, USA

¶ Laboratory of Human Carcinogenesis, National Cancer Institute, Bethesda, Maryland 20892, USA

THE *p53* gene has been a constant source of fascination since its discovery nearly a decade ago^{1,2}. Originally considered to be an oncogene, several convergent lines of research have indicated that the wild-type gene product actually functions as a tumour suppressor gene³⁻⁹. For example, expression of the neoplastic phenotype is inhibited, rather than promoted, when rat cells are transfected with the murine wild-type *p53* gene together with mutant *p53* genes and/or other oncogenes^{3,4}. Moreover, in human tumours, the short arm of chromosome 17 is often deleted (reviewed in ref. 10). In colorectal cancers, the smallest common region of deletion is centred at 17p13.1 (ref. 9); this region harbours the *p53* gene, and in two tumours examined in detail, the remaining (non-deleted) *p53* alleles were found to contain mutations⁹. This result was provocative because allelic deletion coupled with mutation of the remaining allele is a theoretical hallmark of tumour-suppressor genes¹¹. In the present report, we have attempted to determine the generality of this observation; that is, whether tumours with allelic deletions of chromosome 17p contain mutant *p53* genes in the allele that is retained. Our results suggest that (1) most tumours with such allelic deletions contain *p53* point mutations resulting in amino-acid substitutions, (2) such mutations are not confined to tumours with allelic deletion, but also occur in at least some tumours that have retained both parental 17p alleles, and (3) *p53* gene mutations are clustered in four 'hot-spots' which exactly coincide with the four most highly conserved regions of the gene. These results suggest that *p53* mutations play a role in the development of many common human malignancies.

To search for mutations, two approaches were used, both based on the polymerase chain reaction¹² (PCR) (Fig. 1). For tumour cell lines and for xenografts passaged in athymic nude mice, complementary DNA was generated from messenger RNA using oligo(dT) as a primer. A 1,300 base-pair (bp) fragment including the entire *p53* coding region was generated from the cDNA using PCR, and this fragment was cloned and sequenced

in its entirety. For primary tumours, sufficient RNA was often not available for the first approach, and PCR was used to generate a 2.9-kilobase (kb) fragment from tumour DNA. This was the longest fragment that we could reproducibly amplify from the *p53* locus, and included all of the exons found to contain mutations through the first approach.

Using these approaches, we analysed *p53* sequences of tumours derived from the breast, lung, brain, colon or mesenchyme. Tumours of these types have been previously shown to exhibit frequent deletions of chromosome 17p when studied by restriction-fragment length polymorphism (RFLP) methods¹⁰. To test for allelic deletions, tumour DNA samples were digested with *HinfI* and, following Southern transfer, hybridized sequentially to two probes (p14D6 (ref. 13) and pYNZ22.1 (ref. 14)) detecting variable-number tandem-repeat ('VNTR' or 'minisatellite') sequences. DNA samples from normal tissues exhibited two alleles with at least one of these probes in 29 of 31 different individuals tested. Because of this high degree of polymorphism, allelic loss could be assessed with greater than 95% certainty in cell lines and xenografts, even when corresponding normal tissue was not available for comparison.

Nineteen tumours with allelic deletions of chromosome 17p were selected for sequence analysis. Thirteen of the tumours were found to contain a single missense mutation; two tumours each contained two missense mutations; one tumour contained a frame-shift mutation at codon 293; and no mutation was detected in three tumours (Table 1). The PCR reaction is known to be associated with a relatively high rate of base misincorporation¹², and we confirmed this observation by noting several sequence variants (13 out of 34,000 bp sequenced) in individual clones that were not reproducibly present in other PCR reactions from the same tumour sample. All of the mutations listed in Table 1 were confirmed by performing a second PCR reaction and re-sequencing the products *en masse* as described in the legend to Fig. 2.

Two observations indicated that the nucleotide substitutions described in Table 1 represented somatic mutations. First, none of these presumptive mutations have been observed in the sequences of human *p53* genes derived from normal cells, SV40-transformed fibroblasts, or lymphoblastoid cell lines (ref. 15, and references therein). Second, in six cases (tumours 2, 3, 9, 12, 13, 16), normal tissue from the patients whose tumours are described in Table 1 were available for study. To test for the presence of the presumptive mutations in the germline of these patients, a strategy was devised that used both PCR and cloning. Although direct sequencing of PCR products has been shown to be possible by several methods, we found that none of the published methods could be reproducibly applied to all parts of the *p53* coding region. To circumvent this difficulty, we cloned the PCR products into a phagemid vector and used the DNA pooled from 10³ to 10⁴ independent phage clones as a template for DNA sequencing (see legend to Fig. 2). This procedure resulted in sequence data quality as high as that produced using individual plasmid DNA clones as templates, and was used to demonstrate that in each of the six cases noted above, the mutations in the tumour DNA were not present in the germline of the patient (examples in Fig. 2).

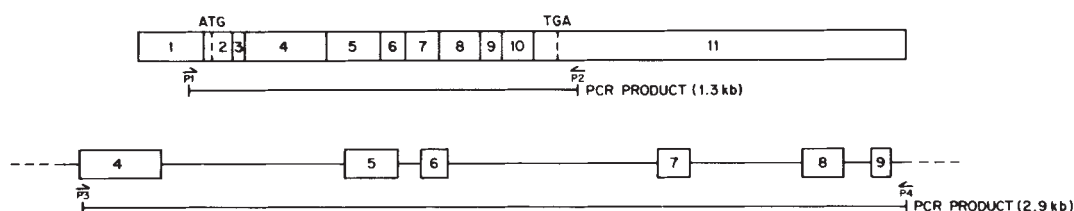
The data described above indicated that most tumours with one 17p allele contained a mutation of the *p53* gene in the remaining allele. To begin to assess the status of tumours that had not lost a 17p allele, we examined cDNA clones from three such tumours. In each case, two cDNA clones derived from PCR products, generated as described in the legend to Fig. 1, were sequenced. In one case (tumour 11), both clones contained a single point mutation at codon 134 (Table 1). In the second case (tumour 16), one clone contained a point mutation at codon 281 and one clone was wild-type. In the third case (tumour 17), both clones were wild-type. To assess the relative expression levels of the mutant alleles, the sequencing strategy described

To whom correspondence should be addressed.

FIG. 1 Strategies for amplification of *p53* gene sequences. Messenger RNA was used to generate a cDNA template for a polymerase chain reaction (PCR) using primers P1 and P2 (top). The PCR product was 1.3 kb long and included the entire coding region.

Alternatively, total genomic DNA was used in a PCR reaction using primers P3 and P4. The PCR product was 2.9 kb long and included exons 4–9 (bottom). The numbered boxes indicate exons and the vertical dotted lines indicate the start (ATG) and stop (TGA) codons respectively.

METHODS. RNA was purified using guanidium isothiocyanate and mRNA selected by binding to Messenger Affinity Paper (Amersham). Complementary DNA was synthesized from 500–750 ng of mRNA using oligo(dT) as a primer. The oligo(dT) primer was removed by isopropanol precipitation; 10 µg of transfer RNA and sodium perchlorate (to a final aqueous concentration of 0.5 M) were added to the reaction, and this was followed by addition of 1/2 volume of isopropanol²⁷. The cDNA was pelleted by centrifugation for 15 min at room temperature and used in a 50-µl PCR reaction consisting



of 35 cycles of 93 °C (1 min), 58 °C (1 min), and 70 °C (2 min). Genomic DNA (2 µg) was used in a 200-µl PCR reaction consisting of 30 cycles at 95 °C (1 min), 58 °C (1 min), and 70 °C (4 min). PCR reactions contained magnesium chloride at a final concentration of 2 mM. The primers used were P1, 5'-GGAATCCACGACGGTGACACG-3'; P2, 5'-GGAATCAAAATGGCAGGGG-AGGG-3'; P3, 5'-GTAGGAATTCGTCCCAAGCAATGGATGAT-3'; P4, 5'-CATCGA-ATTCTGGAAACTTCCACTTGAT-3'. All primers had extraneous nucleotides comprising *Eco*RI sites at their 5' ends to facilitate cloning. The PCR products were digested with *Eco*RI, fractionated by electrophoresis, and following purification from agarose, ligated to *Eco*RI digested Bluescript vectors (Stratagene). Individual clones were sequenced with primers derived from the *p53* coding and intron sequences¹⁵ using T7 polymerase.

in the legend to Fig. 2 was used using cDNA from tumour mRNA as a template. In tumour 11, only the mutant allele was expressed (data not shown); in tumour 16, the mutant and wild-type alleles were expressed at approximately equal levels (Fig. 2, panel 4).

The fact that three (tumours 5, 8 and 10) of the 19 tumours with single chromosome 17p alleles did not contain *p53* gene mutations has potentially important practical and conceptual implications. At least three explanations could account for this result. First, it is possible that mutations existed in these three tumours but were outside the region sequenced. Such mutations could affect the transcription of *p53* mRNA, its stability or translational capacity. Experiments to test this possibility are in progress, but it is notable that although tumour 5 expressed *p53* mRNA, no *p53* protein could be detected on western blots (S. E. Kern and J.M.N., unpublished data). Alternatively, it is possible that a second tumour suppressor gene on chromosome 17p exists and was the target of allelic deletion in some tumours. Finally, the possibility that some allelic deletions represent 'nonselected losses', coincidentally occurring with other independent mutations elsewhere in the genome, cannot be excluded.

Neurofibrosarcomas are tumours that predominantly occur in patients with neurofibromatosis type I. This syndrome can be inherited in an autosomal dominant fashion, and the gene responsible for it (*NF*) has been mapped to chromosome 17q near the centromere¹⁶. Interestingly, chromosome 17 sequences are lost from neurofibrosarcomas from these patients, but the region of deletion sometimes includes the short arm and does not include the region of 17q harbouring the *NF* gene (T. G. and B. Seizinger, unpublished data). The result shown in Table 1 (tumour 12) suggests that a target of allelic deletion, in at least one neurofibrosarcoma, was *p53*.

Altogether, 20 point mutations (19 missense, 1 frameshift) were identified in the present study. These are mapped in Fig. 4, together with the two human *p53* gene missense mutations previously described⁹. Several features are notable. Although the sample size is limited, the mutations tended to be clustered in four hotspots which accounted for 86% of the 21 missense mutations (five mutations in region A, codons 132–143; five mutations in region B, codons 174–179; three mutations in region C, codons 236–248; five mutations in region D, codons 272–281). There have been two missense mutations identified in murine tumour cells, both in the carcinogen-induced fibrosarcoma cell line Meth A: one allele contained a mutation in region A, and the other contained one mutation in region C and one mutation in region D^{17,18}. Interestingly, the four hotspots for *in vivo* mutation coincided exactly with the four most highly conserved

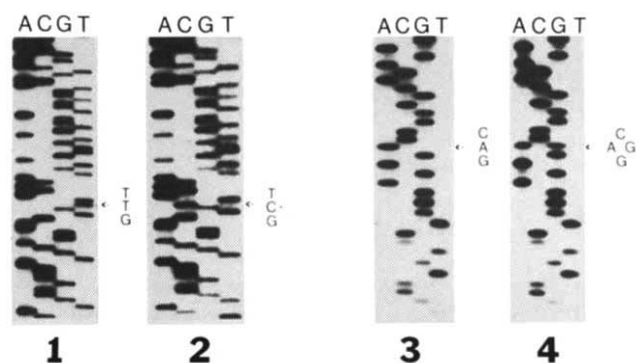


FIG. 2 Examples of sequencing reactions demonstrating *p53* gene mutations. The templates used for the sequencing reactions shown in panels 1–4 consisted of pools of greater than 10³ clones generated from PCR products. Tumour 13 genomic DNA contained a mutation at codon 239 (antisense GCT, panel 2), instead of the wild-type sequence (GTT) found in the genomic DNA from normal lymphocytes from the same patient (panel 1). Panel 4 shows a sequencing reaction of pooled cDNA clones from tumour 16 showing that both wild-type codon 281 (GAC) and mutant codon 281 (GGC) were both expressed. Only the wild-type sequence (GAC) was found in pooled genomic DNA clones from normal lymphocytes of this patient (panel 3).

METHODS. PCR reactions were carried out as described in the legend to Fig. 1, and the reaction products digested with *Eco*RI. The entire reaction was ligated to 0.25 µg of λ ZAP phage vector arms (Stratagene) and packaged using 1/4 of a GLGA-Pack extract (Stratagene). *Escherichia coli* BB4 cells were then infected, and 10³–10⁴ phage clones plated on a 7-cm Petri dish. The λ ZAP vector contains the sequences for a phagemid into which the PCR inserts were cloned, and single-stranded DNA phage can be rescued from the λ phage clones using a helper phage²⁸. An overnight culture of XL-1 Blue cells (Stratagene) was grown in 0.4% maltose and resuspended in 1.5 volumes of 10 mM magnesium sulphate. Phages were eluted from the 7-cm dish in 5 ml phage-dilution buffer (100 mM sodium chloride, 10 mM magnesium sulphate, 20 mM Tris, pH 7.5, 0.02% gelatin) for 2 h at room temperature with gentle agitation. Fifty microlitres of eluate was used to infect 200 µl of XL-1 Blue cells (Stratagene) in the presence of 1 µl helper phage R408 (10¹¹ P.F.U. per ml). After 15 min at 37 °C, 5 ml of 2 × YT broth was added and the culture shaken for 3 h at 37 °C, then heated to 70 °C for 20 min. Cell debris was pelleted at 3,000g for 5 min, and 10 µl supernatant, containing single-stranded DNA phage, was used to infect 200 µl of XL-1 Blue cells, prepared as described above. After 15 min at 37 °C, 100 µl of the mixture (containing over 10⁴ clones determined by titration on XL-1 Blue cells) was inoculated into 50 ml L-broth and shaken overnight at 37 °C. Double-stranded DNA was isolated by alkaline lysis and sequenced as described in the legend to Fig. 1. The primer used for sequencing in panels 1 and 2 was 5'-GAGGCAAGCAGAGGCTGG-3'. The primer used for sequencing in panels 3 and 4 was 5'-TGGTAATCTACTGGGACG-3'.

TABLE 1 *p53* gene mutations in human tumours

Tumour	Tumour name	Tumour type*	Tumour cells tested†	Number of 17p alleles‡	Codon	Mutation Nucleotide	Amino acid
1	D263	Brain	B, X	1	175	CGC → CAC	Arg → His
2	D274	Brain	X	1	273	CGT → TGT	Arg → Cys
3	D303	Brain	B, X	1	216	GTG → ATG	Val → Met
4	D317	Brain	B, X	1	272	GTG → ATG	Val → Met
5	D247	Brain	C	1		None detected	
6	MDA 468	Breast	C	1	273	CGT → CAT	Arg → His
7	T47D	Breast	C	1	194	CTT → TTT	Leu → Phe
8	BT123	Breast	B	1		None detected	
9	1012	Lung	B	1	293	Deleted a G	Frameshift
10	5855	Lung	B	1		None detected	
11	H231	Lung	C	2	134	TTT → TTA	Phe → Leu
12	88-3/14	NFS	B, C	1	179	CAT → TAT	His → Tyr
13	C × 4A	Colon	B, X	1	239	AAC → AGC	Asn → Ser
14	C × 5A	Colon	X	1	248	CGG → TGG	Arg → Trp
15	C × 6A	Colon	X	1	132	AAG → AAC	Lys → Asn
					133	ATG → TTG	Met → Leu
16	C × 7A	Colon	B, X	2	281	GAC → GGC	Asp → Gly
17	C × 19A	Colon	X	2		None detected	
18	C × 20A	Colon	B, X	1	175	CGC → CAC	Arg → His
19	C × 22A	Colon	X	1	175	CGC → CAC	Arg → His
20	C × 26A	Colon	X	1	141	TGC → TAC	Cys → Tyr
21	SW480	Colon	C	1	273	CGT → CAT	Arg → His
					309	CCC → TCC	Pro → Ser
22	SW837	Colon	C	1	248	CGG → TGG	Arg → Trp

* The brain tumours were glioblastoma multiforme; the colon and breast tumours were adenocarcinomas, the NFS tumour was a neurofibrosarcoma developing in a patient with type-I neurofibromatosis; H231 was a small cell carcinoma of the lung, and the other two lung tumours were non-small-cell carcinomas.

† B, Tumour biopsy; C, cell line passaged *in vitro*; X, xenograft derived from biopsy, passaged in athymic nude mice. Whenever two sources of tumour cells are listed, both contained the indicated mutation.

‡ The number of alleles was determined by RFLP analysis, as described in the text.

regions of the *p53* gene, previously identified by Soussi *et al*¹⁹. Of the 41 amino acids contained within regions A-D, 93% are identical in the wild-type *p53* genes of amphibian, avian and mammalian species, compared to a conservation of only 51–57% over the entire *p53* coding sequence. The clustering of mutations and evolutionary conservation of regions A-D suggest that they play a particularly important role in mediating the normal function of the *p53* gene product.

Previous cytogenetic and RFLP studies have shown that allelic deletions of chromosome 17p occur in at least 60% of tumours of the colon, breast, lung, ovaries, cervix, adrenal cortex, bone and bladder, and in at least 30% of brain tumours^{10,20–23}. These tumour types account for most of the neoplasms occurring in humans. Our data suggest that the great majority of tumours with 17p allelic deletions contain mutant *p53* genes, and similar mutations occur in at least some tumours without 17p allelic deletions.

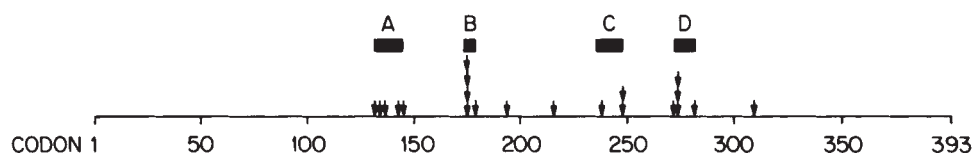
These data lead to the following hypothesis. Mutations in the *p53* gene occur during the process of tumorigenesis, and through a dominant negative effect^{3–5,9,24}, cause tumour progression. The dominant negative effect may be mediated by binding of the mutant *p53* product to the wild-type product, creating an inactive oligomeric complex^{18,25}. Because wild-type products remain in the cell, however, a further loss of growth control can be exerted when the wild-type allele is deleted, leaving the cell with only

a mutant allele. The first example of an intermediate step in this scheme is provided by tumour 16, which had not lost a chromosome 17p allele, but had developed a *p53* gene mutation and expressed both the wild-type and mutant alleles (Table 1 and Fig. 2). We imagine that tumour 16, had it not been surgically removed, would have eventually lost the wild-type *p53* gene through allelic deletion. Indeed, it has been shown that the loss of a chromosome 17p allele is significantly associated with tumour progression in the human host²⁶. Although this model requires much further study before it can be verified, it is now supported by several lines of research, including the demonstration of mutant *p53* genes in most tumours with 17p allelic losses. □

Received 30 August; accepted 30 October 1989.

1. Lane, D. P. & Crawford, L. V. *Nature* **278**, 261–263 (1979).
2. Linzer, D. I. H. & Levine, A. J. *Cell* **17**, 43–52 (1979).
3. Finlay, C. A., Hinds, P. W. & Levine, A. J. *Cell* **57**, 1083–1093 (1989).
4. Ellyahu, D. *et al. Proc. natn. Acad. Sci. U.S.A.* (in the press).
5. Sturzbecher, H.-W., Addison, C. & Jenkins, J. R. *Molec. cell. Biol.* **8**, 3740–3747 (1988).
6. Munroe, D. G., Rovinski, B., Bernstein, A. & Benichimol, S. *Oncogene* **2**, 621–624 (1988).
7. Wolf, D. & Rotter, V. *Proc. natn. Acad. Sci. U.S.A.* **82**, 790–794 (1985).
8. Masuda, H., Miller, C., Koeffler, H. P., Battifora, H. & Cline, M. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7716–7719 (1987).
9. Baker, S. J. *et al. Science* **244**, 217–221 (1989).
10. Cavenee, W. K., Hastie, N. D. & Stanbridge, E. J. eds *Current Communications in Molecular Biology: Recessive Oncogenes and Tumor Suppression* (Cold Spring Harbor Press, New York, 1989).

FIG. 3 Summary of *p53* point mutations in human cancer. Each of the missense mutations listed in Table 1 is indicated with an arrow. In addition, the two point mutations described previously⁹ in human cancers (at codons 143 and 175) are also included. The four regions containing most (86%) of the mutations are indicated by the black bars marked A–D.



11. Knudson, A. *Cancer Res.* **45**, 1437-1443 (1985).
12. Saiki, R. K. et al. *Science* **239**, 487-491 (1988).
13. Kondoleon, S. et al. *Nucleic Acids Res.* **15**, 10605 (1987).
14. Nakamura, Y. et al. *Nucleic Acids Res.* **16**, 5707 (1988).
15. Buchman, V. L. et al. *Gene* **70**, 245-252 (1988).
16. Collins, F. S., Ponder, B. A. J., Seizinger, B. R. & Epstein, C. J. *Am. J. hum. Genet.* **44**, 1-5 (1989).
17. Finlay, C. A. et al. *Molec. cell. Biol.* **8**, 531-539 (1988).
18. Eliyahu, D. et al. *Oncogene* **3**, 313-321 (1988).
19. Soussi, T. et al. *Oncogene* **1**, 71-78 (1987).
20. Lee, J. H. et al. *Proc. Am. Ass. Cancer Res.* **30**, 442 (1989).
21. Atkin, N. B. & Baker, M. C. *Cancer Genet. Cytogenet.* **37**, 229-233 (1989).
22. Yano, T. et al. *J. natn. Cancer Inst.* **81**, 518-523 (1989).
23. Tsai, Y. C. et al. *Cancer Res.* (in the press).
24. Herskowitz, I. *Nature* **329**, 219-222 (1987).
25. Kraiss, S., Quaiser, A., Oren, M. & Montenach, M. *J. Virol.* **62**, 4737-4744 (1988).
26. Kern, S. E. et al. *J. Am. Med. Ass.* **261**, 3099-3103 (1989).
27. Haymerle, H. et al. *Nucleic Acids Res.* **14**, 8615-8624 (1986).
28. Short, J. M., Fernandez, J. M., Sorge, J. A. & Huse, W. D. *Nucleic Acids Res.* **16**, 7583-7600 (1988).

ACKNOWLEDGEMENTS. We thank Y. Nakamura, R. White and M. Litt for providing the RFLP probes and T. Gwiazda for preparation of the manuscript. The work was supported in part by The Clayton Fund, the McAshan Fund, The National Neurofibromatosis Foundation and the NIH.

Putative transcription activator with alternative isoforms encoded by human *ZFX* gene

Ansbert Schneider-Gädicke, Peggy Beer-Romero, Laura G. Brown, Graeme Mardon, Shih-Wen Luoh & David C. Page

Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA, and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

THE *ZFY* gene in the sex-determining region of the human Y chromosome encodes a protein with 13 zinc fingers, may determine whether an embryo develops as a male or female¹. *ZFX*, a related gene on the human X chromosome, may also function in sex determination; it encodes a protein with a very similar zinc-finger domain and escapes X inactivation^{1,2}. *ZFY* and *ZFX* diverged from a common ancestral gene before the radiation of placental mammals, and retain a similar genomic organization². Analysis of complementary DNAs from the mouse Y-chromosomal homologues of *ZFY* indicates that these genes encode probable transcription activators^{3,4}. Here, we report that *ZFX* encodes a protein composed of a highly acidic amino-terminal domain, a basic putative nuclear-localization signal, and a carboxy-terminal zinc-finger domain. This combination of features, also found in the *ZFY* gene product, is typical of transcription activators. Alternative splicing generates *ZFX* transcripts encoding isoforms of 575 and 804 amino acids. These *ZFX* protein isoforms differ in the length of their acidic domains and may be functionally distinct.

ZFX is transcribed in all human cells analysed². We cloned *ZFX* cDNAs from a male lymphoblastoid cell line. Analysis by restriction mapping and hybridization with genomic probes revealed three distinct types of cDNAs. Whereas two types were represented by a single clone each (cDNAs 1 and 3), a third type was represented by three clones (cDNAs 2, 4 and 5).

We determined the nucleotide sequence of one cDNA of each type (Fig. 1). Comparison of cDNAs 1, 2, and 3 showed that alternative splicing and polyadenylation had produced structurally distinct 5' untranslated, coding, and 3' untranslated regions (Fig. 2a). Differential splicing involved an invariant donor site (nucleotide -378) and alternative acceptor sites (at nucleotides +647, -28 and -377, in cDNAs 1, 2 and 3, respectively). Each cDNA contains a single long open reading frame (ORF), and in each case the first ATG occurs in a sequence context favourable for initiation of translation⁵. Complementary

FIG. 1 Composite nucleotide sequence of *ZFX* cDNAs 1, 2 and 3 (Fig. 2a) and predicted amino-acid sequences. Numbering of nucleotides and amino acids is with reference to the first in-frame ATG codon in cDNAs 2 and 3. Known splice sites are indicated by upward arrows. Alternative starts of poly(A) tails are indicated by downward arrows. Complementary DNA 1 begins at nucleotide -527, uses a splice donor site at -378 and an acceptor site at +647; a tail of 20 adenosines follows nucleotide 2,585. Complementary DNA 2 starts at nucleotide -426, uses the same donor, but a different acceptor at -28; a tail of 46 adenosine residues follows nucleotide 5,450. Complementary DNA 3 starts at nucleotide -426, uses the same donor, but yet another different acceptor at -377; a tail of 26 adenosines follows nucleotide 3,298. Complementary DNAs 2 and 3 contain the same ORF with the first in-frame ATG codon at nucleotide 1; the predicted protein of 804 amino acids is composed of an acidic N-terminal domain, a small basic domain (residues 391-406, boxed), and 13 zinc fingers (cysteines of the Cys-X-X-Cys repeats are circled). Complementary DNA 1 has a shorter ORF with the first in-frame ATG codon at nucleotide +688, encoding a protein of 575 residues. The 3' UTR of cDNA 2 contains an *Alu* sequence¹⁶ (nucleotides +3,959 to +4,250, boxed). An AATAAA polyadenylation signal¹⁷ occurs 20 nucleotides 5' of the poly(A) tail of cDNA 2. Similar sequences are present 5' of the poly(A) tails in cDNAs 1 and 3.

METHODS. Complementary DNA libraries were prepared³ using poly(A)⁺ RNA prepared² from human male lymphoblastoid cell line WHT1659. Five million recombinant phages from unamplified libraries were screened using plasmids pDP1007, pDP1041 and pDP1006 (*ZFY* genomic fragments which cross-hybridize to *ZFX*)^{1,2} as probes. Complementary DNA inserts of seven phages were subcloned into Bluescript vectors (Stratagene). The full DNA sequences of three inserts and partial sequences of the other four inserts were determined². As judged by comparison with genomic sequences, five cDNAs originated from *ZFX* and two from *ZFY*.

DNAs 2 and 3, with identical ORFs, encode a protein of 804 amino acids (*ZFX*⁸⁰⁴), whereas cDNA 1 encodes an isoform of 575 amino acids (*ZFX*⁵⁷⁵). (Alternatively, with cDNA 1, translation initiation at a second ATG, whose context is highly favourable, would result in production of an isoform of 573 amino acids.) *In vitro* transcription and translation of cDNAs 3 and 1 yielded the predicted full-length and truncated proteins, respectively (data not shown). Both *ZFX* protein isoforms contain three domains—an N-terminal acidic portion (25% aspartic and glutamic acid), a small basic domain, and a C-terminal run of 13 zinc fingers, each with two cysteines and two histidines (Cys-Cys/His-His zinc fingers). The isoforms differ in that the acidic domain of *ZFX*⁵⁷⁵ is half that of *ZFX*⁸⁰⁴.

Comparison of *ZFX* cDNAs with the genomic locus by restriction mapping and oligonucleotide hybridization gave an overview of the intron-exon organization (Fig. 2a). The C-terminal zinc-finger domain and 3' untranslated region (UTR) are encoded by a single exon, whereas the N-terminal acidic domain is encoded by a minimum of four exons for *ZFX*⁵⁷⁵ (cDNA 1) and a minimum of six exons for *ZFX*⁸⁰⁴ (cDNAs 2 and 3). The cDNAs span 67 kilobases (kb) in the genome.

Of the three types of cDNAs, type 2, encoding *ZFX*⁸⁰⁴, seems to be most representative of the 6.3- and 8-kb transcripts observed on northern blots². First, the coding exon defined by oligonucleotides '150' and '637' is present in cDNAs 2 and 3 but not in cDNA 1 (Fig. 2a). Northern analysis using the corresponding genomic DNA fragment revealed that this exon is present in the 6.3- and 8-kb transcripts (not shown). Second, northern analysis indicated that the polyadenylation site used in cDNA 2 corresponds, at least roughly, to the 3' end of the main *ZFX* transcripts (data not shown). Clone 2, which is 5.6 kb long, could represent a 6.3- or 8-kb transcript that is incomplete at the 5' end.

Transcripts corresponding to cDNA 1, encoding *ZFX*⁵⁷⁵, have not been detected by northern analysis. Using polymerase chain reaction (PCR) amplification, however, we confirmed the differential splicing predicted from cDNA analysis and crucial to the generation of *ZFX* protein isoforms (Fig. 2b). We designed splice-specific 5' primers spanning splices of the invariant donor site at nucleotide -378 to alternative acceptor sites at nucleotides +647, -28 and -377, as in cDNAs 1, 2 and 3, respectively. The

[illegible]

3' primer was common to all three cDNAs. Bulk cDNA generated by reverse transcription of male or female lymphoblastoid RNA was used as template. Fragments predicted from the sequences of cDNA clones 1, 2 and 3 were amplified (Fig. 2b), confirming the differential splicing shown in Fig. 2a. Some unanticipated fragments were simultaneously amplified (Fig. 2b).

The ZFX⁸⁰⁴ protein has a primary sequence strikingly similar to those encoded by the mouse ZFY homologues (Zfy-1 and Zfy-2)^{3,4}, with which it shares virtually identical N and C termini. ZFX⁸⁰⁴ and ZFX⁵⁷⁵ proteins are also similar to the human ZFY protein (ref. 1; and P.B.-R. and A.S.-G., unpublished results). Each of these proteins is composed of an acidic domain, a small basic domain, and 13 Cys-Cys/His-His zinc fingers. This combination of features is reminiscent of those of eukaryotic transcription activators⁶. By analogy, we assume that ZFX and ZFY proteins activate transcription in a sequence-specific fashion. When fused to the DNA-binding domain of GAL4, the acidic domains of mouse Zfx and Zfy-2 do activate transcription in yeast⁷.

Are the various ZFX and ZFY proteins functionally distinct? All share a small basic domain, which could localize the proteins to the nucleus^{8,9}. Cys-Cys/His-His zinc-finger domains^{10,11} bind to nucleic acids in a sequence-specific fashion. The zinc-finger domains of ZFX⁸⁰⁴ and ZFX⁵⁷⁵ are identical and differ from the zinc-finger domain of human ZFY at only 10 of 393 amino acids². Thus ZFX⁸⁰⁴, ZFX⁵⁷⁵ and ZFY proteins could all recog-

nize the same nucleic-acid target sequences, and thereby perhaps regulate transcription of the same genes.

The N-terminal portions of the ZFX and ZFY proteins are highly acidic. Acidic domains mediate activation in several eukaryotic transcription factors¹²⁻¹⁵. The potency of such activating domains may depend more on the number and density of acidic residues than on particular primary sequences⁵. In ZFX⁸⁰⁴ and mouse Zfx, Zfy-1 and Zfy-2, the acidic domains are of similar size, each with ~25% of residues acidic. By contrast, ZFX⁵⁷⁵ contains an abbreviated acidic domain. ZFX⁸⁰⁴ and ZFX⁵⁷⁵ might be expected to have quantitatively different activating potencies. Furthermore, acidic domains of transcription factors are believed to interact with other proteins in transcription complexes. Truncation of the acidic domain in ZFX⁵⁷⁵ could alter its interactions with other factors, resulting in qualitatively different regulatory properties. (Internal ATG codons in human ZFY and mouse Zfx, Zfy-1 and Zfy-2 could allow production of similarly truncated isoforms.)

We compared the human ZFX cDNA sequence with that of a murine Zfx cDNA (ref. 7; Fig. 3). The coding regions of ZFX and Zfx show a high degree of similarity. It is of interest that the 3' UTR is also conserved. Except for an *Alu* insertion in the human sequence, the entire human 3' UTR of 3 kb shows over 80% nucleotide identity with the murine 3' UTR.

Given the similarity in the primary structure of the ZFX protein and its Y chromosome-encoded homologues, it is conceivable that the products of the ZFX and ZFY genes are

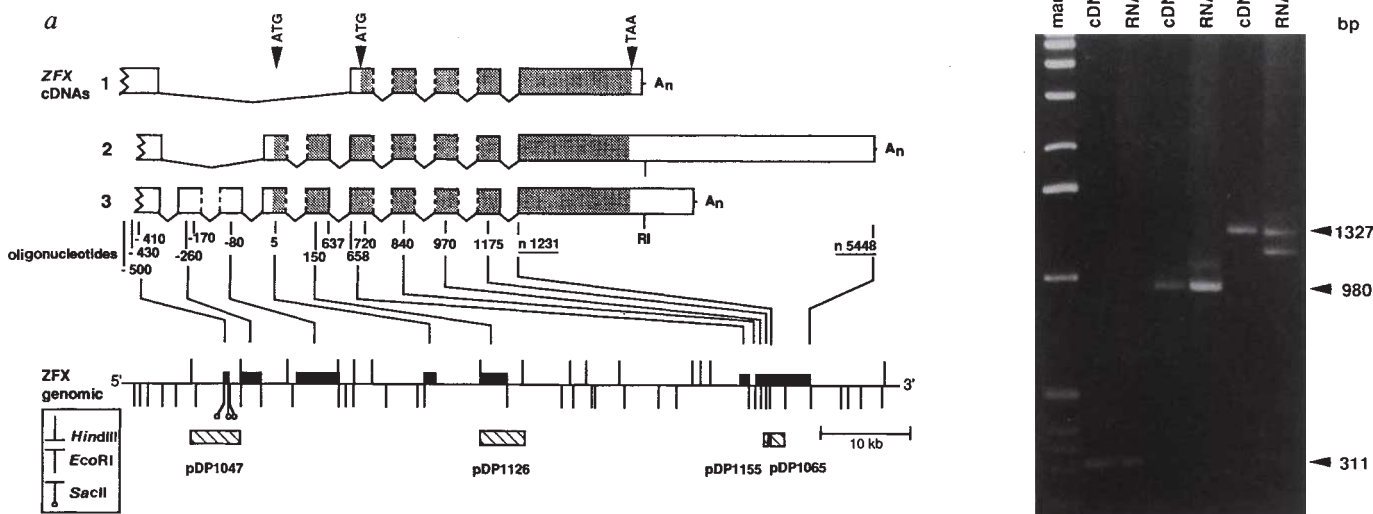
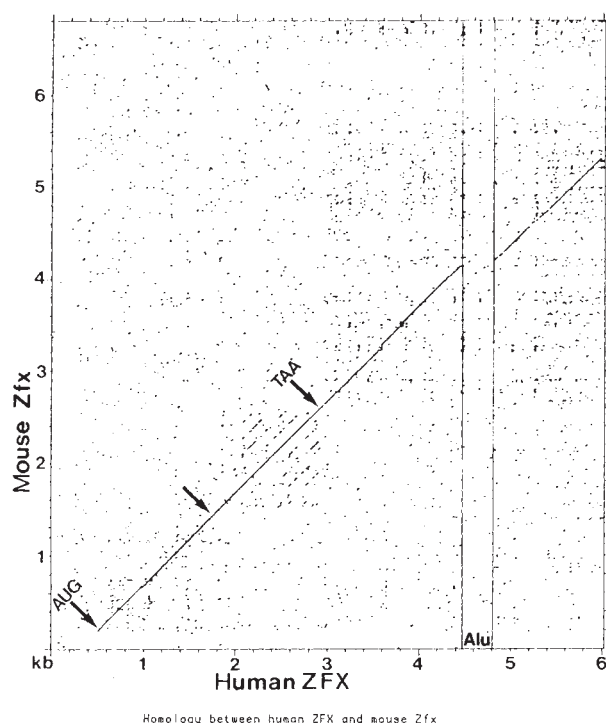


FIG. 2 Analysis of alternative ZFX transcripts. *a*, Organization of cDNAs and genomic locus. Top: boxes indicate the minimum number of exons and, with the exception of the 3'-most exon, are not to scale; coding regions are shaded. Several splice sites (solid vertical lines of boxes) were defined by DNA sequence comparison among cDNA clones 1, 2 and 3 or by comparison with genomic sequences. Additional splices (dashed vertical lines of boxes) are not precisely defined but inferred from comparison of cDNA and genomic restriction maps. Boxes bounded by dashed lines may contain more than one exon each. Genomic DNA fragments² hybridizing (5–10 °C below calculated melting temperature) with ZFX oligonucleotides (numbers denote the 11th nucleotide in 20-mer oligonucleotides) are shown by black segments; each contains one or more exons. *b*, PCR analysis of differential splicing. Three different 5' primers (20mers) spanning splices unique to cDNAs 1, 2 or 3 have in common the invariant donor site at nucleotide –378, but each includes a unique acceptor (–389 to –378/+647 to +654 (primer 5'1); –383 to –378/–28 to –15 (primer 5'2); –381 to –362 (primer 5'3)). An invariant 3' primer (+923 to +946) was used. Using cDNA clones 1, 2 and 3 as templates, the predicted products were amplified (cDNA 1, 311 base pairs (bp); cDNA 2, 980 bp; cDNA 3, 1,327 bp). Using 'RNA' (bulk cDNA

prepared from male lymphoblastoid WHT1659 RNA by reverse transcription) as template, products of the same lengths were obtained. In addition, with bulk cDNA, primer 5'2 amplified a 1.1-kb fragment (which may incorporate additional coding exon(s) not detected by cDNA cloning), whereas primer 5'3 amplified a barely detectable 1.4-kb fragment (possibly incorporating additional coding or 5' UTR exon(s)) and a 1.2-kb fragment (which may lack coding or 5' UTR exon(s)) present in cDNA 3). Controls not shown: (1) The 5' primers are splice specific, yielding products only with the corresponding cDNAs; (2) amplification of cell line RNA, without reverse transcriptase, yielded no products; (3) by using bulk cDNA from female lymphoblastoid cell line WHT1628 as template, products of the same lengths were obtained; and (4) hybridization of Southern blots of these gels verified that the ethidium bromide-detected products originate from ZFX.

METHODS. Poly(A)⁺ RNA (200 ng) was reverse-transcribed using Moloney reverse transcriptase (Bethesda Research). Forty cycles of amplification were performed using *Taq* polymerase and a DNA Thermo Cycler (Cetus). Each PCR reaction (10%) was separated on an agarose gel and stained with ethidium bromide. The marker lane contains a '1-kb ladder' (Bethesda Research).



Homology between human ZFX and mouse Zfx

	% Identity	No of gaps
DNA	protein	
Coding region:		
N-terminal domain	90 92	2
Zinc finger domain	93 99.5	0
3' untranslated region:		
5' of Alu insertion	85 -	19
3' of Alu insertion	81 -	22

FIG. 3 Nucleotide and corresponding amino-acid similarity between human *ZFX* (Fig. 1) and murine *Zfx* (ref. 7). Striking similarity is found not only in the coding region, but also across the entire 3' UTR (except for an *Alu* insertion, indicated by vertical lines, in the human), and ~20 nucleotides 5' of the ATG codon. Arrows indicate the start codon in *ZFX* cDNAs 2 and 3, the splice site preceding the zinc-finger exon, and the TAA stop codon. Numbering (in kb) is with reference to the 5' ends of the cDNAs. The analysis¹⁸ employed a 'window' of 21 and a 'stringency' of 14.

functionally interchangeable. But the occurrence of isoforms—at least of *ZFX*—indicates that *ZFX* and *ZFY* could produce transcription factors with qualitatively or quantitatively distinct functions. The 3' UTRs of *ZFX* transcripts are longer and more highly conserved than those found in *Zfy* transcripts^{3,4}, and could provide a basis for differential post-transcriptional regulation. Detailed studies of *ZFX* and *ZFY* transcripts during development should further elucidate their postulated roles in sex determination. □

Received 7 August; accepted 17 October 1989.

- Page, D. C. *et al.* *Cell* **51**, 1091–1104 (1987).
- Schneider-Gädick, A. *et al.* *Cell* **57**, 1247–1258 (1989).
- Mardon, G. & Page, D. C. *Cell* **56**, 765–770 (1989).
- Ashworth, A., Swift, S. & Affara, N. *Nucleic Acids Res.* **17**, 2864 (1989).
- Kozak, M. *Cell* **44**, 283–292 (1986).
- Ptashe, M. *Nature* **335**, 683–689 (1988).
- Mardon, G. *et al.* *Molec. cell. Biol.* (in the press).
- Dang, C. V. & Lee, W. M. F. *Molec. cell. Biol.* **8**, 4048–4054 (1988).
- Dingwall, C. & Laskey, R. A. *Rev. Cell Biol.* **2**, 367–390 (1986).
- Miller, J., McLachlan, A. D. & Klug, A. *EMBO J.* **4**, 1609–1614 (1985).
- Brown, R. S., Sander, C. & Argos, P. *FEBS Lett.* **186**, 271–274 (1985).
- Ma, J. & Ptashe, M. *Cell* **48**, 847–853 (1987).
- Hope, I. A. & Struhl, K. *Cell* **46**, 885–894 (1986).
- Hollenberg, S. M. & Evans, R. M. *Cell* **55**, 899–906 (1988).
- Godowski, P. J., Picard, D. & Yamamoto, K. R. *Science* **241**, 812–816 (1988).
- Schmid, C. W. & Jelinek, W. R. *Science* **216**, 1065–1070 (1982).
- Fitzgerald, M. & Shenk, T. *Cell* **24**, 251–260 (1981).
- Maizel, J. V. & Lenk, R. P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 7665–7669 (1981).

ACKNOWLEDGEMENTS. We thank Nazneen Aziz, Cornelius Murre and Sang-Ho Park for technical advice, and our colleagues, especially Eric Lander, for comments on the manuscript. This work was supported by the NIH, the Whitaker Health Sciences Fund, the Lucille P. Markey Charitable Trust, and the Searle Scholars Program/Chicago Community Trust. A.S.G. was supported by a fellowship from EMO.

PDGF induction of tyrosine phosphorylation of GTPase activating protein

Christopher J. Molloy, Donald P. Bottaro, Timothy P. Fleming, Mark S. Marshall*, Jackson B. Gibbs* & Stuart A. Aaronson

Laboratory of Cellular and Molecular Biology, National Cancer Institute, Building 37, Room 1E24, Bethesda, Maryland 20892, USA

* Department of Molecular Biology, Merck Sharp and Dohme Research Laboratories, West Point, Pennsylvania 19486, USA

THE cascade of biochemical events triggered by growth factors and their receptors is central to understanding normal cell-growth regulation and its subversion in cancer. Ras proteins (p21^{ras}) have been implicated in signal transduction pathways used by several growth factors, including platelet-derived growth factor (PDGF)¹. These guanine nucleotide-binding Ras proteins specifically interact with a cellular GTPase-activating protein (GAP)². Here we report that in intact quiescent fibroblasts, both AA and BB homodimers of PDGF rapidly induce tyrosine phosphorylation of GAP under conditions in which insulin and basic fibroblast growth factor (bFGF) are ineffective. Although GAP is located predominantly in the cytosol, most tyrosine-phosphorylated GAP is associated with the cell membrane, the site of p21^{ras} biological activity^{3,4}. These results provide a direct biochemical link between activated PDGF-receptor tyrosine kinases and the p21^{ras}–GAP mitogenic signalling system.

In human malignancies, *ras* genes are commonly activated as oncogenes and their p21 products are potently mitogenic when microinjected into quiescent fibroblasts^{5,6}. Yet the function of p21^{ras} proteins in normal mitogenic signal transduction remains unknown. It is well established that the GTP-bound forms of p21^{ras} proteins are active and that the GDP-bound forms are inactive^{7–9}. The GAP protein of relative molecular mass 125,000 (*M_r* ~ 125K) catalyses the conversion of p21-GTP to p21-GDP so that the rate of this recycling is more than 100-fold faster than that intrinsic for p21 (ref. 2). GAP has no effect on oncogenic p21 proteins, allowing them to remain in their active GTP-bound states.

To investigate whether GAP serves as a substrate for activated growth factor-receptor kinases in intact cells, confluent NIH3T3 cells were rendered quiescent by incubating them overnight in serum-free medium. Cells were exposed to saturating concentrations of PDGF BB homodimer for 5 min, after which cell lysates were subjected to immunoblot analysis with an antiphosphotyrosine antiserum (anti-P-Tyr) or anti-GAP peptide serum 638 (ref. 10). PDGF BB homodimer induced the specific tyrosine phosphorylation of several proteins (Fig. 1a). These included the α - and β -PDGF receptors each of *M_r* 185K (ref. 11), a protein of *M_r* 145K that has been identified as phospholipase C γ (ref. 12), and several other proteins of *M_r* ranging from 50–125K. Anti-GAP peptide serum 638, raised against GAP amino-acid residues 139–152 (ref. 10), identified a single band corresponding to a protein of *M_r* ~ 125K, which was not observed in the presence of excess competing peptide (Fig. 1a, lane 3).

When quiescent cells or cells treated with PDGF BB homodimer were first immunoprecipitated with either polyclonal or monoclonal anti-P-Tyr antibodies followed by immunoblotting with anti-GAP antiserum 638, GAP was detectable only in cells stimulated with PDGF BB homodimer (Fig. 1b, lane 3). Similar results were observed using anti-GAP peptide serum 677 raised against GAP residues 968–981 (ref. 10). Moreover, recovery of GAP by anti-P-Tyr antiserum was specifically inhibited by an excess of competing phosphotyrosine

(Fig. 1b, lane 4), but not by phosphoserine or phosphothreonine (data not shown). Whereas these experiments were performed using saturating amounts of PDGF BB homodimer (100 ng ml^{-1}), increased levels of anti-P-Tyr-recoverable GAP were also observed using concentrations of PDGF BB that did not maximally induce DNA synthesis (data not shown). Taken together, these results imply that GAP is an early substrate of PDGF-receptor tyrosine kinases or is specifically co-immunoprecipitated with a tyrosine-phosphorylated substrate(s).

When cell lysates were first immunoprecipitated with an anti-serum against recombinant GAP, a protein of M_r 125K was detected from either unstimulated (Fig. 1c, lane 1) or PDGF BB homodimer-stimulated (Fig. 1c, lane 2) cells. Note that we reproducibly observed slightly better recovery of GAP from PDGF-stimulated cell lysates (Fig. 1c, lane 2). When parallel immunoprecipitates were probed with anti-P-Tyr antiserum, only the 125K GAP band extracted from PDGF-stimulated cells was identified (Fig. 1c, lane 4). These results confirm that GAP was tyrosine-phosphorylated in response to PDGF BB homodimer. The GAP protein contains at least two consensus tyrosine-phosphorylation sites at residues 161 and 720 (ref. 13). Recent studies have also indicated that GAP can serve as a substrate of tyrosine kinases *in vitro*^{14,15}. The precise location

of the *in vivo* sites of GAP tyrosine phosphorylation remains to be determined.

We next monitored the recovery of GAP by anti-P-Tyr antiserum at different times after PDGF BB homodimer stimulation of quiescent cells. PDGF BB homodimer induced rapid PDGF-receptor autophosphorylation, which increased over the first 10 min and then gradually decreased (Fig. 2a). Similarly, increased recovery of GAP by anti-P-Tyr was observed within 1 min of PDGF addition, was maximal by ~10 min, and persisted for at least 120 min (Fig. 2b). In other experiments, we detected increased levels of anti-P-Tyr-recoverable GAP for up to 8 h after PDGF BB homodimer stimulation (data not shown).

GAP has previously been reported to be located in the cell cytosol². Because we found that GAP was rapidly tyrosine-phosphorylated in response to PDGF, we performed subcellular fractionation studies to determine the site of GAP tyrosine phosphorylation within the cell. In quiescent fibroblasts, GAP was found predominantly in the cytosolic fraction (Fig. 3a, lane 2). But after stimulation of the cells with PDGF BB homodimer for 10 min, we observed a modest but reproducible increase in GAP associated with the particulate fraction concomitant with a slight decrease in the cytosolic component (Fig. 3a, lanes 3 and 4). Anti-P-Tyr antiserum-recoverable GAP was observed only in the PDGF BB homodimer-stimulated cells and was most abundant in the particulate fraction (Fig. 3b, lane 3). By comparing the relative signal intensities of anti-GAP immunoblots of anti-P-Tyr-recovered versus total GAP in the respective subcellular fractions, we estimated that at least 15–20% of the membrane-associated GAP protein was recovered by anti-P-Tyr antiserum immunoprecipitation and that <2% of cytosolic GAP could be recovered under the same conditions. Because anti-P-Tyr immunoprecipitation is relatively inefficient (C.J.M., unpublished results), it is probable that a much greater fraction of the particulate (P-100)-localized GAP is tyrosine-phosphorylated in response to PDGF BB homodimer.

The specificity of PDGF-induced GAP phosphorylation was analysed by comparing the responses of quiescent NIH3T3 cells to stimulation with saturating concentrations of several growth factors. The addition of either PDGF AA homodimer, bFGF, or insulin induced specific tyrosine phosphorylation of cellular proteins (Fig. 4a). For example, bFGF caused tyrosine phosphorylation of a previously identified protein of M_r 90K (ref. 16), and insulin induced tyrosine phosphorylation of proteins

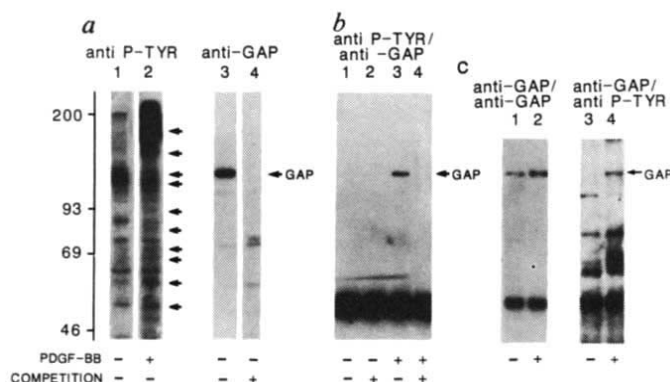


FIG. 1 Tyrosine phosphorylation of GAP in PDGF BB homodimer-stimulated NIH3T3 cells. **a**, Lanes 1 and 2, immunoblot of total cell lysates ($100 \mu\text{g}$) using anti-P-Tyr antiserum. Cells were either quiescent or treated with PDGF BB homodimer for 5 min as indicated. Lanes 3 and 4, immunoblot of cell lysates, using anti-GAP peptide serum 638 (ref. 10) in the absence (lane 3) or presence (lane 4) of competing peptide. **b**, Immunoblot of anti-P-Tyr antiserum-recovered proteins using anti-GAP antiserum. Where noted, phosphotyrosine (5 mM) was added to the lysates before immunoprecipitation for competition. Lanes 1 and 2, anti-P-Tyr antiserum immunoprecipitates from quiescent cells; lanes 3 and 4, anti-P-Tyr antiserum immunoprecipitates from PDGF BB homodimer-stimulated cells. **c**, Anti-GAP serum 106 immunoprecipitates from quiescent (lanes 1 and 3) and PDGF BB homodimer-stimulated (lanes 2 and 4) cells were analysed by immunoblotting using either anti-GAP peptide serum 638 or anti-P-Tyr antiserum as indicated. **METHODS.** After incubation in serum-free medium for 18 h, confluent NIH3T3 cells were either left untreated or stimulated with PDGF BB homodimer (Amgen) at 100 ng ml^{-1} for 5 min. Cultures were rinsed twice in ice-cold PBS containing 1 mM Na_2VO_4 , and lysed on ice in 0.5 ml P-Tyr lysis buffer (50 mM HEPES buffer pH 7.5, 1% Triton X-100, 50 mM NaCl, 50 mM NaF, 10 mM sodium pyrophosphate, 5 mM EDTA, 1 mM Na_2VO_4 , 1 mM PMSF, in addition to $10 \mu\text{g ml}^{-1}$ aprotinin, leupeptin, and pepstatin A). Lysates were clarified by centrifugation at $12,000g$ at 4°C for 10 min, and then immunoprecipitated with affinity-purified anti-P-Tyr antiserum^{11,19}, monoclonal anti-P-Tyr antibodies (PY 20, PY 69, ICN), or rabbit anti-GAP serum 106 raised against recombinant GAP (residues 128–1,044; J.B.G., unpublished results) for 2 h on ice. Immune complexes were recovered using protein G-agarose (Gammabind G, Genex). Immunoprecipitates were washed five times in P-Tyr lysis buffer and then solubilized in SDS-PAGE sample buffer as previously described^{19,29}. Aliquots of the whole-cell lysates were removed before immunoprecipitation for immunoblot analysis. Gel electrophoresis and immunoblotting were performed as previously described^{11,29}. Immunoreactive bands were visualized using ^{125}I -labelled protein A followed by autoradiography.

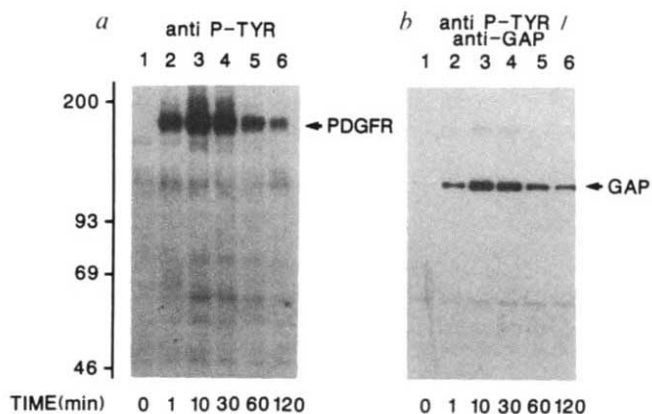


FIG. 2 Time course of anti-P-Tyr antiserum recovery of GAP after PDGF BB homodimer stimulation. **a**, Anti-P-Tyr antiserum immunoblot analysis of whole-cell lysates from unstimulated cells (lane 1) and cells stimulated by PDGF BB homodimer for the indicated time (lanes 2–6). Arrow denotes the autophosphorylation of PDGF receptors¹¹. **b**, Anti-GAP peptide serum 677 immunoblot analysis of anti-P-Tyr-antiserum immunoprecipitates from cells either unstimulated (lane 1) or stimulated with PDGF BB homodimer for the indicated times. Arrow denotes the immunoreactive GAP protein recovered by anti-P-Tyr antiserum after PDGF BB homodimer stimulation. Methods for immunoprecipitation and immunoblotting are described in the legend to Fig. 1. Blots were exposed for 15 h using an intensifying screen.

with M_r s of 175K and 95K, the latter corresponding to the β -subunit of the insulin receptor^{17,18}. Stimulation with both PDGF AA and BB homodimers led to increased anti-P-Tyr recovery of GAP (Fig. 4b, lanes 2 and 3). In other studies using NIH3T3 cells overexpressing epidermal growth factor receptors¹⁹, rapid anti-P-Tyr antiserum recovery of GAP has also been observed after stimulation with epidermal growth factor (C.J.M., unpublished data). By contrast, neither insulin nor bFGF, under conditions in which bFGF was as mitogenically active as the PDGF homodimers, was able to induce detectable GAP tyrosine phosphorylation (Fig. 4b, lanes 4 and 5). Thus there seem to be differences in the efficiency of GAP tyrosine phosphorylation in response to different growth factors.

These studies demonstrate that GAP is tyrosine phosphorylated *in vivo* in response to specific growth factors. Recent evidence indicates that PDGF and epidermal growth factor receptor kinases also rapidly induce tyrosine phosphorylation of other proteins, including the *c-ras* gene product²⁰, phosphatidylinositol kinase²¹, and phospholipase C_γ (refs 22–24). The rapid and sustained time course of anti-P-Tyr antiserum recovery of GAP that we observed is very similar to that recently reported for PDGF-stimulated tyrosine phosphorylation of phospholipase C_γ (ref. 24). Furthermore, as reported for phospholipase C_γ (ref. 24), our preliminary experiments indicate that GAP co-immunoprecipitates with activated PDGF receptors (C.J.M., unpublished data). Taken together, these findings

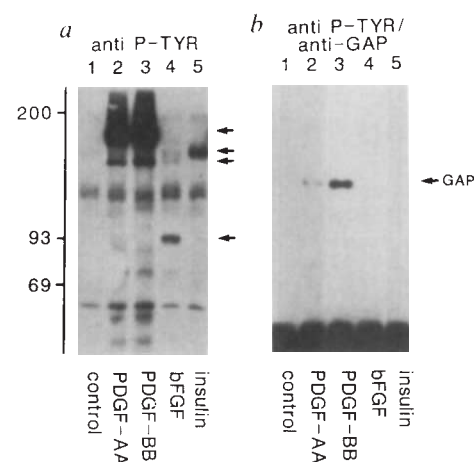


FIG. 4 Anti-P-Tyr antiserum recovery of GAP in response to different growth factors. *a*, Anti-P-Tyr antiserum immunoblot analysis of whole-cell lysates from unstimulated (lane 1) cells or cells treated for 10 min with PDGF AA homodimer (100 ng ml⁻¹; lane 2), PDGF BB homodimer (100 ng ml⁻¹; lane 3), basic FGF (200 ng ml⁻¹; lane 4), or insulin (100 μ g ml⁻¹; lane 5). Tyrosine-phosphorylated bands induced by different growth factors are indicated by arrows. *b*, Immunoblot of anti-P-Tyr antiserum-recovered proteins using anti-GAP peptide serum. Anti-P-Tyr antiserum immunoprecipitates were prepared from cells treated as in *a*. Methods are as described in Fig. 1. Blots were exposed to film for 4 days.

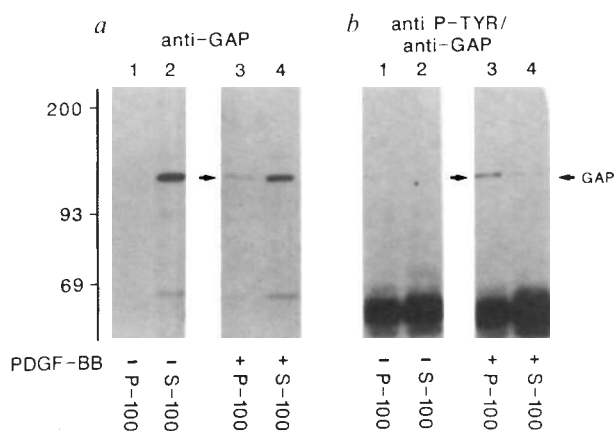


FIG. 3 Subcellular localization of GAP following PDGF BB homodimer stimulation. *a*, Anti-GAP peptide serum 638 immunoblot of protein extracts (40 μ g per lane) from either particulate (P-100) or cytosolic (S-100) cell fractions. Cells were either quiescent (lanes 1 and 2) or stimulated with PDGF BB homodimer for 10 min (lanes 3 and 4). *b*, Anti-GAP peptide serum 638 immunoblot of anti-P-Tyr antiserum-recovered proteins from subcellular fractions of unstimulated (lanes 1 and 2) and PDGF BB homodimer-stimulated (lanes 3 and 4) cells. Blots were exposed to film for 5 days.

METHODS. After incubation in serum-free medium for 18 h, confluent NIH3T3 cells (3×100 -mm plates) were either left untreated or stimulated with PDGF BB homodimer at 100 ng ml⁻¹ for 10 min. Cultures were then rinsed twice with cold PBS containing 1 mM Na₃VO₄, and then 0.75 ml hypotonic lysis buffer (20 mM HEPES buffer pH 7.4, 5 mM sodium pyrophosphate, 5 mM EGTA, 1 mM MgCl₂, 1 mM Na₃VO₄, 1 mM PMSF, in addition to 10 μ g ml⁻¹ aprotinin, leupeptin, and pepstatin) was added to each plate on ice. Cells were collected by scraping into a Dounce homogenizer with a tight-fitting pestle and subsequently lysed with 40 strokes. After centrifugation at 1000g for 5 min, the supernatant fraction was ultracentrifuged at 100,000g for 45 min. The cytosolic (S-100) supernatant fraction was removed, and the insoluble pellet was resuspended and recentrifuged. The resulting pellet (particulate fraction) was then solubilized in 1 ml P-Tyr lysis buffer (see legend to Fig. 1). Cytosolic fractions were concentrated to 1 ml by ultrafiltration using Centricon-10 filters (Amicon). Under these conditions, cell protein recovery in P-100 and S-100 fractions were about equal. Aliquots (40 μ g) from each fraction were used for anti-GAP peptide serum immunoblot analysis. For phosphotyrosine analysis, aliquots of each fraction containing 400 μ g protein were first immunoprecipitated with anti-P-Tyr antiserum followed by immunoblotting with anti-GAP peptide serum as described in Fig. 1 legend.

argue that GAP is an early direct substrate of certain growth factor receptors.

Tyrosine phosphorylation is known to modulate the activity of several enzymes, including the *src* gene product^{25,26}. It is also well established that the p21^{ras} molecule must be located in the membrane to be biologically active^{1,3}. Thus, our demonstration that GAP recovered by anti-P-Tyr antiserum is associated predominantly with the particulate fraction further indicates that tyrosine phosphorylation may alter the functional interaction of GAP and p21^{ras}. In one proposed model, GAP serves as an attenuator of p21^{ras} function². If so, tyrosine phosphorylation of GAP could transiently remove its inhibition, allowing the p21^{ras}-GTP complex to effect mitogenic signalling through some unknown downstream components. Another model proposes that GAP can act as a downstream effector as well as attenuator of p21^{ras} function^{2,27,28}. Thus, tyrosine phosphorylation of GAP could modulate its effector role, perhaps by coupling the GAP-p21^{ras} complex to other molecules. By whatever mechanism tyrosine phosphorylation perturbs GAP-p21^{ras} function, our studies provide direct biochemical evidence linking these important growth regulatory molecules to the mitogenic signalling cascade activated by growth factors. □

Note added in proof. We have also observed tyrosine phosphorylation of GAP in NIH3T3 cells transformed by oncogenes encoding non-membrane-spanning as well as membrane-spanning tyrosine kinases.

Received 4 August; accepted 10 October 1989.

- Mulcahy, L. S., Smith, M. R. & Stacey, D. W. *Nature* **313**, 241–243 (1985).
- Trahey, M. & McCormick, F. *Science* **238**, 542–545 (1987).
- Willingham, M. C., Paston, I., Shih, T. Y. & Scolnick, E. M. *Cell* **19**, 1005–1014 (1980).
- Willumsen, B. M. et al. *Nature* **310**, 583–586 (1984).
- Stacey, D. W. & Kung, H.-F. *Nature* **310**, 508–511 (1984).
- Feramisio, J. R. et al. *Cell* **38**, 109–117 (1984).
- DeVendittis, E., Vitelli, A., Zahn, R. & Fasano, O. *EMBO J.* **5**, 3657–3663 (1986).
- Marshall, M. S., Gibbs, J. B., Scolnick, E. M. & Sigal, I. S. *Molec. cell. Biol.* **7**, 2309–2315 (1987).
- Gibbs, J. B. et al. *J. Biol. Chem.* **262**, 10426–10429 (1987).
- Marshall, M. S. et al. *EMBO J.* **8**, 1105–1110 (1989).
- Matsui, T. et al. *Science* **243**, 800–804 (1989).
- Wahl, M. I. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1568–1572 (1989).
- Vogel, U. S. et al. *Nature* **335**, 90–93 (1988).
- Gibbs, J. B. et al. in *The Guanine-Nucleotide Binding Proteins: Common Structural and Functional Properties* (eds Bosch, L., Kraal, B. & Parmeggiani, A.) 191–200 (Plenum, New York, 1989).
- Gibbs, J. B. et al. in *ADP-Ribosylating Toxins and G-Proteins: Insights into Signal Transduction* (eds Moss, J. & Vaughn, M.) (American Society of Microbiology, Washington DC) (in the press).
- Coughlin, S. R. et al. *J. Biol. Chem.* **263**, 988–993 (1988).
- Pang, D. T. et al. *J. Biol. Chem.* **260**, 7131–7136 (1985).

18. Rosen, O. M. *Science* **237**, 1452-1458 (1987).
19. DiFiore, P. P. *et al. Cell* **51**, 1063-1079 (1987).
20. Morrison, D. K., Kaplan, D. R., Rapp, U. & Roberts, T. M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8855-8859 (1988).
21. Kaplan, D. R. *et al. Cell* **50**, 1021-1029 (1987).
22. Wahl, M. I. *et al. Molec. cell. Biol.* **9**, 2934-2943 (1989).
23. Margolis, B. *et al. Cell* **57**, 1101-1107 (1989).
24. Meisenhelder, J., Suh, P.-G., Rhee, S. G. & Hunter, T. *Cell* **57**, 1109-1122 (1989).
25. Hunter, T. & Cooper, J. A. *Rev. Biochem.* **54**, 897-930 (1985).
26. Pivnicka-Worms, H. *et al. Cell* **49**, 75-82 (1987).
27. Cales, C., Hancock, J. F., Marshall, C. J. & Hall, A. *Nature* **332**, 548-550 (1988).
28. McCormick, F. *Cell* **56**, 5-8 (1989).
29. DiMarco, E. *et al. Oncogene* **4**, 831-838 (1989).

ACKNOWLEDGEMENTS. We thank A. Ng for preparation of rabbit anti-GAP serum 106, and J. Thompson for photography.

Sequence-specific RNA binding by the HIV-1 Rev protein

Maria L. Zapp & Michael R. Green

Department of Biochemistry and Molecular Biology, Harvard University, 7 Divinity Avenue, Cambridge, Massachusetts 02138, USA

THE human immunodeficiency virus type 1 (HIV-1) Rev protein acts post-transcriptionally to increase the amounts of the viral *gag-pol* and *env* messenger RNAs in the cytoplasm of infected cells^{1,2}. The mechanism of Rev action is uncertain. Possibilities include an accelerating effect on the rate of export of its mRNA targets from the nucleus³⁻⁶ and/or modulation of the splicing of pre-mRNAs (reviewed in ref. 7). Both the *gag-pol* and *env* mRNAs contain a sequence that is required for responsiveness to Rev—the Rev responsive element, RRE^{3,4,6,8}. Here we show that Rev is a sequence-specific binding protein, whose binding site is the RRE. This information should help to clarify the mechanism by which Rev acts.

RNA binding experiments were performed using extracts or fractions containing or lacking Rev expressed in *Escherichia coli* from plasmid pZ76-Rev. This plasmid was constructed by inserting a *rev* complementary DNA clone into the expression vector plasmid pRW76, introduced into *E. coli* strain XA-90, and *rev* expression induced by treatment with isopropylthiogalactoside (IPTG). In XA-90 (pZ76-Rev) cells, IPTG specifically induced the expression of a polypeptide of relative molecular mass 18,000 (M_r 18K), which was not observed after IPTG treatment of non-transformed XA-90 cells (Fig. 1a). An immunoblot analysis using a polyclonal α -Rev antibody confirmed that the IPTG-induced 18K polypeptide was Rev (Fig. 1b).

An RNase protection-mobility shift assay was designed to test whether Rev binds to RNA. The RNA binding substrate was an SP6-derived ³²P-labelled RNA containing sequences from the HIV-1 *Bgl*II-*Bam*HI fragment (positions 7,201-7,721; ref. 9), which encompasses the RRE (Fig. 2, bottom; pGEM-HIV1 RRE). After incubation of the ³²P-labelled RNA substrate in the *E. coli* extracts, the reaction mixtures were extensively treated with RNases T1 and A, and ³²P-labelled RNA-protein complexes resolved on a native polyacrylamide-agarose gel.

A protein in the Rev-containing extract bound to the RRE-containing ³²P-labelled RNA substrate (Fig. 2). An RNase-resistant ³²P-labelled RNA fragment of retarded electrophoretic mobility was present in the extract prepared from IPTG-treated XA-90 (p76-Rev) cells (lane 4), but not in extracts prepared from non-induced cells (lane 3), or extracts prepared from IPTG-treated XA-90 cells (lane 2). A partially purified Rev fraction (Fig. 1b; DEAE, 0.2) also supported formation of the ³²P-labelled RNA complex (lane 6), whereas complex formation was not detected in a DEAE fraction that lacked Rev (Fig. 1b; DEAE, 0.4; Fig. 2, lane 5). The ³²P-labelled RNA complex was eliminated by addition of SDS (lane 7) or proteinase K (lane 8), indicating that it contained both RNA and protein. Finally,

the ³²P-labelled RNA-protein complex was not formed if the RNase was added before rather than after the addition of the Rev-containing extract (lane 8). Taken together these results indicate that a protein binds to the RRE-containing RNA substrate and protects a region of it from RNase digestion.

Formation of the RNA-protein complex is specific to an RRE-containing substrate. An unrelated adenovirus RNA substrate (MINX, lane 11), or an HIV-1 RNA substrate that contains sequences upstream of the RRE (580, lane 10), did not

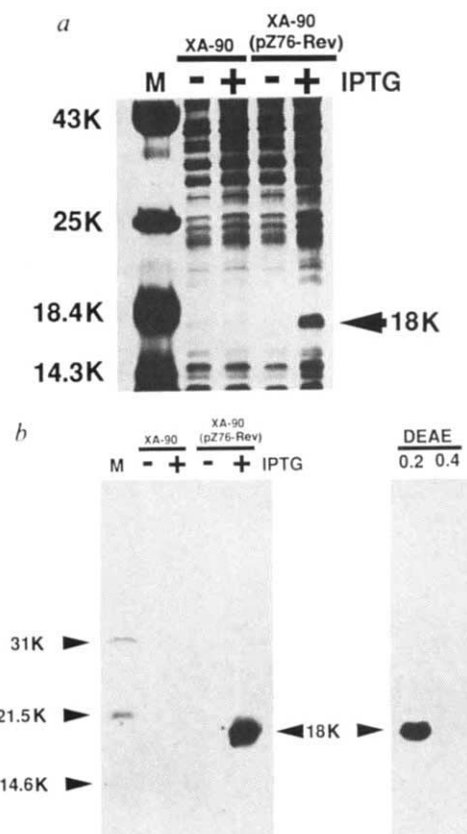


FIG. 1 Detection of Rev synthesized in *E. coli*. *a*, Coomassie blue-stained 12.5% SDS-polyacrylamide gel of protein ($\sim 35 \mu\text{g}$) from the various extracts. Extracts were prepared from either XA-90 or XA-90 (pZ76-Rev) cells with or without IPTG treatment as indicated. Migration of protein M_r standards is shown on the left. The arrow denotes the IPTG-induced polypeptide in XA-90 (pZ76-Rev) cells of $M_r \sim 18\text{K}$. *b*, Immunoblot analysis. Protein ($\sim 3.5 \mu\text{g}$) from the various extracts was fractionated on a 12.5% SDS-polyacrylamide gel, and Rev detected by immunoblotting using a polyclonal α -Rev antibody. The unfractionated extracts (left) were identical to those used in *a*. Chromatographic fractions from a DEAE-Sephacryl column loaded with an IPTG-induced XA-90 (pZ76-Rev) extract are shown on the right. Migration of ¹⁴C-labelled M_r standards (M) is shown on the left.

METHODS. *a*, Plasmid pZ76-Rev was constructed by inserting the *Eco*NI-*Hind*III fragment from the plasmid pcREV (ref. 17), which contains a *rev* cDNA clone, between the *Nco*I and *Hind*III sites of an *E. coli* expression vector plasmid pRW76 (pRW76 is a derivative of the plasmid pLPK76-7; ref. 18). The *Eco*NI site (position 6,005) is within the *rev* cDNA ATG translation initiation codon and was joined to the *Nco*I site of pRW76 to restore the translation initiation site and reading frame. IPTG-induction and preparation of sonicated *E. coli* extracts is described in ref. 19. *b*, The α -Rev antibody was prepared by immunizing rabbits with an affinity-purified *E. coli*-derived function protein containing Rev sequences encoded by plasmid pSVX1 (ref. 4). The α -Rev antibody (1:1,500 dilution) was incubated with the nitrocellulose for 1 h at 4 °C. Antibody-antigen complexes were detected using ¹²⁵I-labelled protein A (NEN) and autoradiography. DEAE fractions were prepared by loading an IPTG-treated XA-90 (pZ76-Rev) extract onto a DEAE-Sephacryl column in buffer A (ref. 19) in 50 mM NaCl. After collection of a flow-through fraction, the column was eluted with steps of 100 mM, 200 mM, 400 mM and 1 M NaCl. Although all fractions contained significant amounts of total protein, immunodetectable Rev was present only in the 200 mM fraction.

support formation of the RNA-protein complex. Likewise, the protein also did not bind to the antisense transcript of plasmid pGEM-HIV1 RRE (α -RRE, lane 12).

To verify that the RNA-binding protein was in fact Rev, three additional experiments were performed (Fig. 3). First, the size of the RNA-binding protein was determined. The proteins in a Rev-containing extract were separated by preparative SDS-PAGE. The gel lane was divided into eight slices, the proteins eluted from each slice, the SDS removed by acetone precipitation, and the precipitated proteins resuspended in 6 M guanidine hydrochloride to denature the polypeptide chains^{10,11}. After removing the guanidine hydrochloride by dialysis, each fraction was assayed for RNA-binding activity. All of the RNA-binding activity was in the region of the SDS-polyacrylamide gel corresponding to an M_r of 18–20K (Fig. 3a), the only fraction that contained Rev (Fig. 1b, and data not shown).

Next, we asked whether Rev was physically associated with the 32 P-labelled RNA substrate. If so, an α -Rev antibody should

bind to the 32 P-labelled RNA-protein complex and alter its electrophoretic mobility. 32 P-labelled RNA-protein complexes were formed as described above. Before loading the complexes on the gel, either the α -Rev antibody or a control pre-immune serum was added to the reaction mixtures (Fig. 3b). As the amounts of the polyclonal α -Rev antibody added increased, the band corresponding to the normal 32 P-labelled RNA-protein complex diminished and concomitantly several bands of reduced electrophoretic mobility appeared. By contrast, the addition of comparable amounts of a control pre-immune serum did not affect either the amount or the electrophoretic mobility of the normal 32 P-labelled RNA-protein complex.

Finally, we investigated whether Rev could be directly detected in the RNA-protein complex. The RNA-protein complex was resolved on a native gel, the band excised, the proteins resolved on a SDS-polyacrylamide gel, and Rev detected by immunoblot analysis. The immunoblot (Fig. 3c) confirmed that the 18K Rev protein was indeed present in the RNA-protein complex.

These experiments show that Rev binds specifically to an RRE-containing RNA substrate. To demonstrate that Rev bound to the RRE, various derivatives of the RNA substrate were analysed (Fig. 4). The 208-nucleotide *StyI*-*Sau3AI* fragment (positions 7,361–7,569) is the minimal sequence that enables a viral RNA to be affected by Rev (ref. 3). Figure 4a shows the results of RNA binding experiments using 3' deletion derivatives of plasmid pGEM3-HIV1 RRE. Rev bound equally well to an RNA whose 3' end was at the *Sau3AI* site (position 7,569) as it did to the original substrate whose 3' end is at the *HindIII* site (position 7,721). But Rev did not bind to an RNA whose 3' end was at the *HinfI* site (position 7,540) or the *AluNI* site (position 7,491). Thus, the 3' boundary of the region required for Rev binding is within the RRE, between positions 7,540 and 7,569.

To determine the 5' boundary of the Rev-binding site, a comparable 5' deletion analysis was performed (Fig. 4b). Rev bound equally well to an RNA whose 5' end was at the *MboII* site (position 7,312) as it did to the original substrate whose 5' end was at the *BglII* site (position 7,201). But further 5' deletion within the minimal RRE to the *HaeIII* site (position 7,424) or *AluNI* site (position 7,491) eliminated Rev binding. Thus, the 5' boundary of the region required for Rev binding is within the RRE, between positions 7,312 and 7,424.

Our *in vitro* binding data is in excellent agreement with the results of a previous *in vivo* study³, which defined a minimal RRE based on Rev function. Deletions that abolish RRE function also prevent Rev binding. Therefore the minimal RRE apparently corresponds to the minimal Rev-binding site. The RNA region required for Rev binding is relatively large (minimally 116 nucleotides, from the *HinfI* site to the *HaeIII* site). Perhaps Rev binding requires an RNA structure that is formed by two or more non-contiguous RNA regions. In fact, the RRE is within an RNA region predicted to have extensive secondary structure³.

Relatively few sequence-specific RNA-binding proteins have been analysed and only one protein motif, the ribonucleoprotein (RNP) consensus sequence (reviewed in ref. 12), has been shown to be involved in sequence-specific RNA binding¹³. Rev does not contain an RNP consensus, indicating that its RNA-binding motif could be of a novel class. The N-terminal portion of Rev is rich in arginine residues and mutations in this region can interfere with Rev function¹⁴. It is tempting to speculate that this arginine-rich basic region contributes to the RNA-binding activity of Rev.

Besides this arginine-rich region, Rev contains at least one other functional domain located near the C terminus¹⁴. Our results imply that one of the functional domains of Rev corresponds to the RNA-binding region (presumably the basic region). The other functional domain(s) presumably provides an effector function, which could interact with a cellular com-

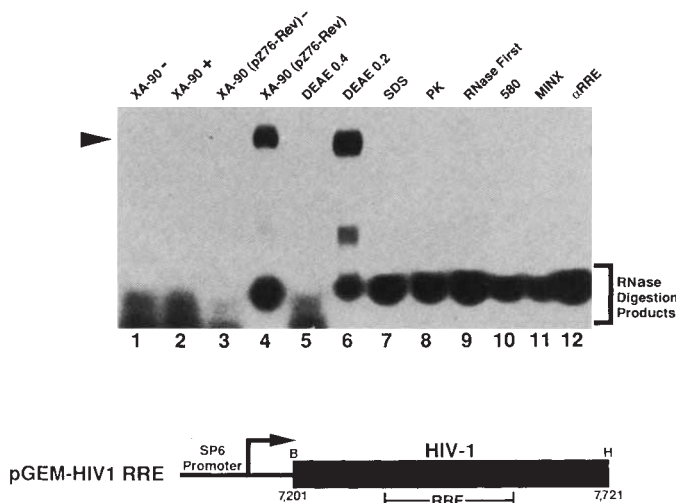
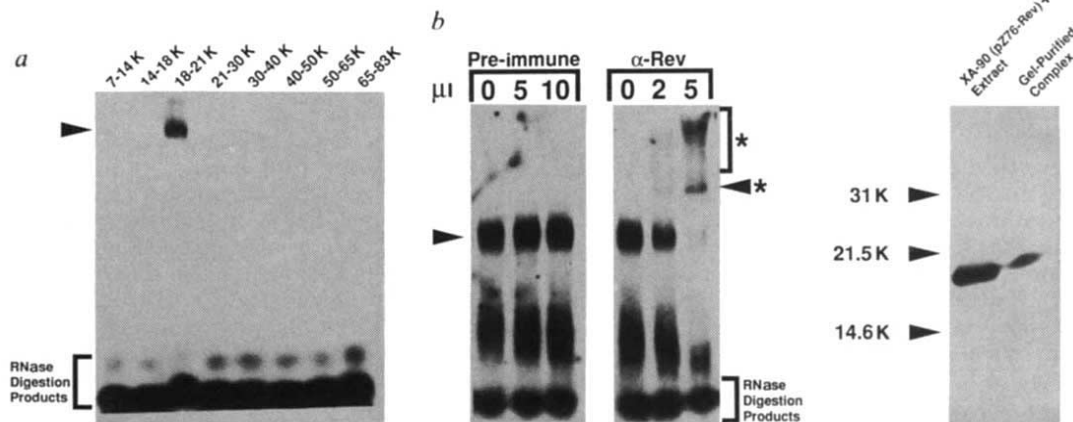


FIG. 2 Analysis of RNA binding by an RNase protection-mobility shift assay. The standard RNA-binding reaction contained the plasmid pGEM-HIV1 RRE 32 P-labelled RNA substrate and extract prepared from IPTG-treated XA-90 (pZ76-Rev) cells (lane 4). The main variation from this standard reaction is indicated above each lane. Lanes 1–9, plasmid pGEM-HIV1 RRE 32 P-labelled RNA substrate. Lanes 4 and 7–12, extract prepared from IPTG-treated XA-90 (pZ76-Rev) cells. Lanes 1 and 2, extract prepared from non-induced and IPTG-induced XA-90 cells, respectively. Lane 3, extract prepared from non-induced XA-90 (pZ76-Rev) cells. Lanes 5 and 6, DEAE (0.4 M) and 0.2 M fractions of an IPTG-treated XA-90 (pZ76-Rev) cell extract, respectively. Lanes 7 and 8, identical to lane 4 except that 0.1% SDS or 0.4 μ g proteinase K, respectively, were added to the sample before loading on the gel. Lane 9, similar to lane 4 except that the 32 P-labelled RNA substrates were incubated with the RNases before addition of the extract. Lane 10, HIV-1 580 RNA substrate. Lane 11, adenovirus MINX RNA substrate²⁰. Lane 12, antisense transcript of plasmid pGEM-HIV1 RRE. The position of the RNA-protein complex is indicated by the arrowhead. The structure of the 32 P-labelled RNA probe run-off at the *HindIII* site (position 7,721) is shown below the autoradiograph. Black box, HIV-1 sequences; B, *BglII* site; H, *HindIII* site. The minimal HIV-1 RRE (positions 7,361–7,569) as defined in ref. 3 is indicated.

METHODS. 32 P-labelled RNA probes were synthesized by *in vitro* transcription using SP6 or T7 RNA polymerase and full-length transcripts purified on a 5% denaturing polyacrylamide gel. RNA binding reactions contained 10–20 μ l extract or fraction, 2 μ l 32 P-labelled RNA probe (36,000 Cerenkov c.p.m.), and buffer A (ref. 19) in a final volume of 25 μ l. Binding reactions were at 30 °C for 15 min, after which 2.5 U RNase T1 (Pharmacia) and 1.25 μ g RNase A (Pharmacia) were added and incubation continued for 10 min. The reaction mixtures were directly loaded onto a native 6% polyacrylamide (60:1 cross-linking ratio)–0.5% agarose gel. The electrophoresis buffer was 100 mM Tris-HCl, pH 8.8, 50 mM glycine. The 580 RNA substrate was synthesized from a plasmid in which the 580-nucleotide *BglII*–*BglII* HIV-1 DNA fragment (positions 6,621–7,200) was cloned into the *Bam*HI site of the plasmid pGEM3 (Promega) polylinker in the orientation such that the sense strand was transcribed using SP6 RNA polymerase.

FIG. 3 Rev is the RNA-binding protein. *a*, Guanidine renaturation of RNA-binding activity. Each RNA-binding reaction contained 20 μ l of the gel-purified protein fraction indicated above the lane.

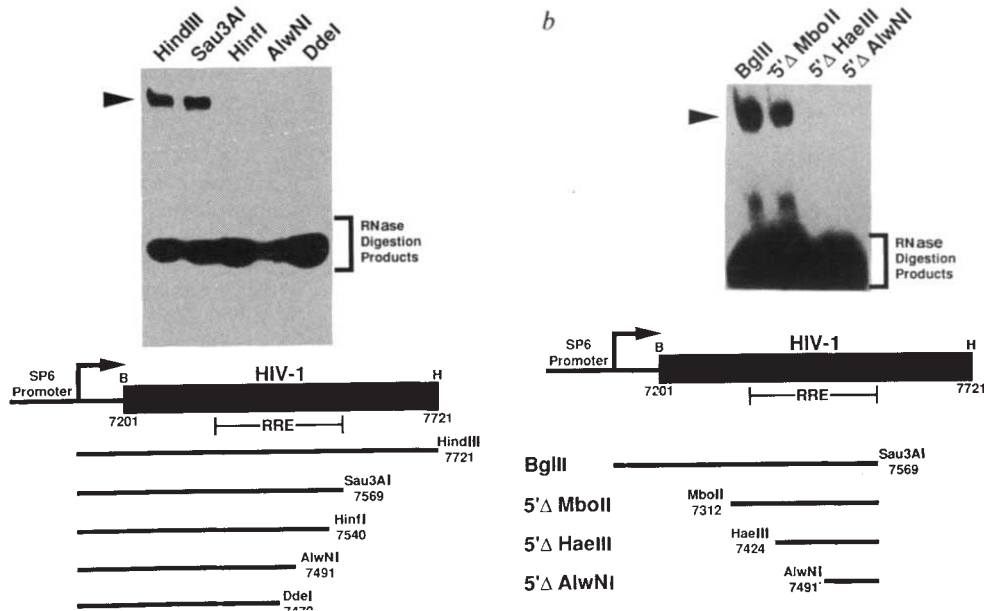
METHODS. Extract (2.3 mg) prepared from IPTG-treated XA-90 (pZ76-Rev) cells was fractionated on a preparative 12.5% SDS-polyacrylamide gel. The gel lane was divided into eight slices, each slice was diced using a razor blade, and soaked overnight at room temperature in 1 ml elution buffer¹⁰. Polyacrylamide was removed by centrifugation, proteins precipitated by addition of five volumes of 100% acetone and incubation on dry ice for 3 h. Precipitated proteins were collected by centrifugation, rinsed in 80% acetone-20% elution buffer without SDS, resuspended in 50 μ l of 6 M guanidine hydrochloride-0.1% Nonidet P-40-0.1 M KCl and incubated overnight at room temperature. Samples were extensively dialysed against buffer A (ref. 19). *b*, An α -Rev antibody alters the electrophoretic mobility of the ³²P-labelled RNA-protein complex. The type and amount (μ l) of antiserum added to the reaction mixtures is indicated above each lane. The position of the normal RNA-protein complex is indicated by the arrowhead on the



left. The positions of RNA-protein complexes of altered electrophoretic mobility are indicated by the starred arrowhead and brackets on the right. The faster-migrating complex is due to partial proteolysis of Rev in this particular extract. *c*, Rev is present in the ³²P-labelled RNA-protein complex. The ³²P-labelled RNA-protein complex resolved on a native gel was located by autoradiography, the gel slice excised, and directly inserted into a lane of a 12.5% SDS-polyacrylamide gel. Rev was detected by immunoblot analysis using the α -Rev antibody as described in the legend to Fig. 1*b*.

FIG. 4 Delineation of the Rev-binding site. *a*, 3' deletion analysis. Truncated ³²P-labelled RNA substrates were synthesized from the plasmid pGEM-HIV1 RRE DNA templates cleaved at the indicated restriction sites. Each RNA binding reaction contained an identical amount of ³²P-labelled RNA substrate. A schematic diagram of the ³²P-labelled RNA substrates is shown below the autoradiograph. *b*, 5' deletion analysis. Each RNA-binding reaction contained an identical amount of ³²P-labelled RNA substrate. A schematic diagram of the ³²P-labelled RNA substrates is shown below the autoradiograph.

METHODS. 5' deletion constructs were prepared by inserting either the *Mbo*II-*Hind*III (positions 7,312-7,721), *Hae*III-*Hind*III (positions 7,424-7,721), or *Alw*NI-*Hind*III (positions 7,491-7,721) HIV-1 DNA fragments between the *Hinc*II and *Hind*III sites of plasmid pGEM4 (Promega). SP6-transcription templates were prepared by cleaving these constructs with *Sau*3A1.



ponent involved, for example, in mRNA nuclear export or pre-mRNA splicing. Some mutations within this C-terminal domain have a dominant-negative phenotype¹⁴. A Rev protein that contains an RNA-binding domain but lacks the putative effector function would compete with wild-type Rev for binding to the RRE and thus have a dominant-negative phenotype.

The HTLV1 Rex protein seems to function similarly to Rev (see refs 15, 16). It is therefore likely that Rex is also a sequence-specific RNA-binding protein. Rex can functionally substitute for Rev in co-transfection experiments¹⁵. This raises the possibility that Rev and Rex bind to the same RNA site even though they have no detectable amino-acid similarity. □

Received 27 September; accepted 31 October 1989.

- Feinberg, M. B., Jarrett, R. F., Aldovini, A., Gallo, R. C. & Wong-Staal, F. *Cell* **46**, 807-817 (1986).
- Sodroski, J. *et al.* *Nature* **321**, 412-417 (1986).
- Malim, M. H., Hauber, J., Le, S.-Y., Maizel, J. V. & Cullen, B. R. *Nature* **338**, 254-257 (1989).

- Hammarskjöld, M.-L. *et al.* *J. Virol.* **63**, 1959-1966 (1989).
- Felber, B. K., Hadzopoulou-Cladaras, M., Cladaras, C., Copeland, T. & Pavlakis, G. N. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1495-1499 (1989).
- Emmerman, M., Vazeux, R. & Peden, K. *Cell* **57**, 1155-1165 (1989).
- Green, M. R. & Zapp, M. L. *Nature* **338**, 200-201 (1989).
- Hadzopoulou-Cladaras, M. *et al.* *J. Virol.* **63**, 1265-1274 (1989).
- Ratner, L. *et al.* *Nature* **313**, 277-284 (1985).
- Hager, D. A. & Burgess, R. R. *Analyt. Biochem.* **109**, 79-86 (1980).
- Briggs, M. R., Kadonaga, J. T., Bell, S. P. & Tjian, R. *Science* **234**, 47-51 (1986).
- Bandzilulis, R. J., Swanson, M. S. & Dreyfuss, G. *Genes Dev.* **3**, 431-437 (1989).
- Query, C. C., Bentley, R. C. & Keene, J. D. *Cell* **57**, 89-101 (1989).
- Malim, M. H., Bohnlein, S., Hauber, J. & Cullen, B. R. *Cell* **58**, 205-214 (1989).
- Rimsky, L. *et al.* *Nature* **335**, 738-740 (1988).
- Seiki, M., Inoue, J.-I., Hidaka, M. & Yoshida, M. *Proc. natn. Acad. Sci. U.S.A.* **85**, 7124-7128 (1988).
- Malim, M. H., Hauber, J., Fenrick, R. & Cullen, B. R. *Nature* **335**, 181-183 (1988).
- Keegan, L., Gill, G. & Ptashne, M. *Science* **231**, 699-704 (1986).
- Lin, Y. S., Carey, M. F., Ptashne, M. & Green, M. R. *Cell* **54**, 659-664 (1988).
- Zillmann, M., Rose, S. D. & Berger, S. M. *Molec. cell. Biol.* **7**, 2877-2883 (1987).

ACKNOWLEDGEMENTS. We thank Drs M.-L. Hammarskjöld, D. Rekosh and D. Knight for gifts of α -Rev antibodies; Drs B. Cullen and W. Greene for HIV-1 plasmids; Y. Li and J. Liu for help in the initial phase of this work; and Dr M. Carey for advice on *E. coli* expression. M.L.Z. holds a NIH postdoctoral fellowship. This work was supported by a grant from the NIH (M.R.G.).

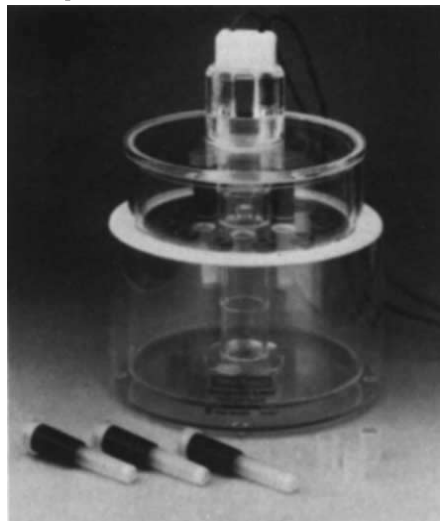
Pre-Christmas product selection

A gel elutor for extracting nucleic acids or protein, a cryopreservation system for freezing/thawing mouse embryos and a neurotoxicological evaluation service based on the leech brain are just a few of the products outlined below.

New from Bioproducts for Science, Inc. is a **mouse anti-rat natural killer cell monoclonal antibody** that recognizes and reacts with all rat NK cells, large granular lymphocytes and neutrophils (*Reader Service No. 100*). The mouse anti-rat NK cell mAb has applications in the study of transplant immunology and in studies requiring *in vivo* depletion of NK cells. The mAb is available in three forms: 5-ml vials of culture supernatant with 5 µg/ml of antibody; 0.25-ml vials of ascites with 1–2 mg/ml of antibody; and as a F(ab')₂ FITC conjugate in 1-ml vials containing 700 µg/ml of antibody. All products are supplied with a procedures guide.

Handy helpers

According to Fisher Scientific, the Fisher Biotech Gel Eluter can simultaneously extract samples from up to six gel slices using a **vertical gel elution technique** (*Reader Service No. 101*). The vertical design of the apparatus is said to offer advantages over conventional horizontal elution methods by accelerating migration of the sample, minimizing electroconvection and reducing build-up around the sieve pores. The elution column is divided

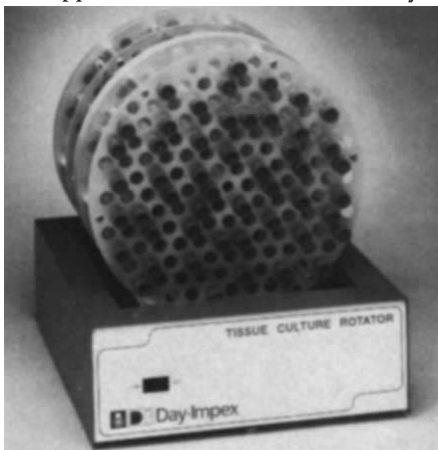


FisherBiotech Gel Eluter: designed for vertical gel elution.

by a porous frit into a sample chamber and removable recovery chamber. The frit acts as a physical barrier to the gel slice while allowing the eluted material to pass into the recovery chamber, to be drawn off by syringe. Fisher says, the eluter will recover virtually any nucleic acid or protein sample with a 75–100 per cent recovery efficiency. The apparatus will work with most power supplies and each elution column requires a current of 1–2 mA.

The \$295 (US) gel eluter can also be used as a tube gel isoelectric focusing apparatus. It can run up to 12 tube gels under identical conditions used in IEF analysis or as the first dimension separation in 2-D electrophoresis.

For the smooth, continuous **rotation of bottles and test tubes**, Day-Impex has introduced a new line of Tissue Culture Rotators (*Reader Service No. 102*). Typical applications include the monolayer



Day-Impex tissue culture rotators provide variable speed and angle of rotation.

culture of anchorage-dependent cells for increased production of, for example, monoclonal antibodies, viruses, enzymes and other secreted products. Four models of the bottle rotator are available which hold between two and ten bottles and operate at speeds of 0.5–8 r.p.m. A tachometer on the control panel provides a readout of rotational speed. The tissue culture test tube rotators are designed to operate in incubators up to temperatures of 60 °C. A single-drum rotator has a capacity of 84 25-mm tubes or 150 16-mm tubes and can operate at rotational speeds of up to 70 r.p.m. The tube capacity can be doubled by the addition of an extra drum. The angle of the drum is adjustable between 5° and 90° above horizontal. All rotator drums are constructed of enamel-coated stainless steel for corrosion resistance. Prices for the bottle and tube rotators range from £480–1,200 (UK).

Save space in your fridge/freezer with the new plastic **square media bottles** from Nalge Co., which are now available in 30- and 60-ml sizes (*Reader Service No. 103*). The PETG square bottles do not break or shatter like glass and have O₂/CO₂ barrier properties that prevent outgassing and destructive pH shifts in the stored media,



Space-saving square media bottles from Nalge.

says Nalge. The transparent bottle and leakproof linerless closure are made of materials that are non-toxic and non-pyrogenic. Aseptic technique is facilitated by the deep closure and narrow-mouth design, and the pouring ring on the neck aids dispensing. Researchers can choose from three different resins for bottles that are autoclavable, chemical and temperature resistant. The 30-ml and 60-ml bottles are \$81.60 and \$105.60 (US), respectively, for a case of 96 bottles.

In a spin

Measuring only 12 cm in diameter, the MC 200 NanoFuge **microcentrifuge** is, says Hoefer, ideal for applications requiring spinning at low *g*-forces (*Reader Service No. 104*). The \$355 (US) NanoFuge can hold up to six 1.5- or 2-ml microcentrifuge tubes or, when fitted with an adaptor, will hold tubes as small as 0.25, 0.4 and



The MC 200: Hoefer's tiny microcentrifuge for spinning at low *g*-forces.

0.5 ml. Stated applications include spinning down droplets from caps and walls of microcentrifuge tubes, blood, urine and tissue centrifugation, and the sedimentation of bacterial cells. Closing the lid of the instrument initiates the spinning process and opening it stops it. The NanoFuge can operate at a maximum speed of 5,000 r.p.m. with a g-force of 200.

Scientific Equipment Design has launched a new generation of **refrigerated centrifuges** that employ solid-state refrigeration technology, eliminating the need for CFCs, bulky compressors or gas refrigeration (*Reader Service No. 105*). The £2,985 (UK) 20R high-speed microcentrifuge can be supplied with a large number of rotor options that will hold anything from 120 microtubes to six 25-ml tubes. The 20R can operate at maximum speeds of 11,000 r.p.m. The £3,796.50 (UK) 30R refrigerated benchtop centrifuge has a capacity of two litres and can operate at speeds of 3,000 r.p.m. or 5,000 r.p.m. if fitted with smaller rotors. Cooling is provided down to 2 °C and the temperature can be set to within 0.1 °C by the digital temperature display/control. According to SED, the bowl temperature should equilibrate within 10 minutes of switch-on. Safety features of the 20R and 30R include an automatic out-of-balance indicator and safety cut-out, dual-lid interlock and three-position brake.

Frozen assets

Transgenic Sciences, Inc. has developed Cryozyte — a **cryoprotectant system** designed for the long-term, low-temperature storage of mouse embryos (*Reader Service No. 106*). The procedure takes about an hour and permits the ultra-rapid cooling of mouse embryos without ice crystallization. Once preserved in Cryo-

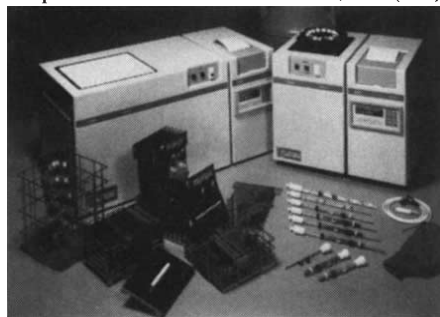


Cryopreserve mouse embryos with Transgenic Sciences' Cryozyte kit.

zyte, TSI says the embryos can be stored indefinitely in any standard liquid nitrogen containers that accommodate canes. Viable embryos are incubated in equilibrium medium and then loaded into plastic straws that are prefilled with vitrification medium. Embryos are recovered by thawing at room temperature, followed by incubation in dilution and rehydration media. The \$235 (US) kit comes complete

with straws, plugs, step-by-step protocol guide, and all the necessary solutions for the cryopreservation of 375 mouse embryos.

Two new **controlled-rate freezer units** — Kryo 10/16 and 10/1.7 — are available from Planar Biomed (*Reader Service No. 107*). The freezers consist of a benchtop controller with a separate and insulated freeze-chamber module which is filled by the programmed transfer of liquid nitrogen. The Kryo 10/16, costing \$11,500 (US) for the basic unit, has a 16-litre capacity and is suitable for freezing blood bags, ampoules and straws. The \$11,750 (US)



Controlled-rate freezers from Planar Biomed. Kryo 10/1.7 model has a smaller capacity but straws and ampoules can be placed both horizontally and vertically. The temperature range in both units is +40 °C to -180 °C with cooling rates variable between 0.01 °C per minute to -50 °C per minute. The microprocessor-based programmer monitors the chamber and sample temperature with platinum-resistance thermometers: any malfunctioning is indicated on the display. Up to ten sets of program parameters can be stored.

Jencons Scientific Ltd are the sole distributors in the UK for the Taylor-Wharton **cryostorage systems** — the K Series — where K denotes the storage capacity in thousands of 2-ml vials (*Reader Service No. 108*). The K Series consists of 3K, 8K, 17K and 27K units, although a 5K unit will be available in early 1990. The refrigerator units can be provided with dedicated inventory systems for both liquid- and vapour-phase sample storage. Automatic control options — alarm only, autofill or full automatic control — for monitoring the level of refrigerant are available for the 8K, 17K and 27K models. The modular design of the K Series allows several units to be connected up to a single liquid nitrogen source, while each can remain under independent control.

DNA extraction

Using **glass adhesion methods**, extract DNA from agarose, separate out unincorporated nucleotides, remove organic solvents, isolate DNA from minipreps and concentrate DNA with the USBioclean kit from USB (*Reader Service No. 109*). Based on the methods of Vogelstein and

Gillespie, 2 µl of the glass powder will typically bind 1 µg of DNA with 80 per cent recovery of the starting material, says USB. The \$80.50 (US) USBioclean kit comprises glass powder suspension, NaI solution, concentrated rinse buffer, Tris-Acetate-EDTA electrophoresis buffer and pipette tips. It is unsuitable for DNA extractions from acrylamide gels, but can be used for both low- and high-melt agarose. DNA that has been extracted by this method, which, USB says, takes about 20–25 minutes, can be used for restriction digests and other enzymatic manipulation.

Perkin-Elmer Cetus is offering the Isogene kit for the **extraction and purification of DNA** (*Reader Service No. 110*). The method uses a DNA binder, which, the company says, has a higher capacity than the glass-milk variety — typically 5 µg of DNA is bound using 10 µl of binder. Recoveries of bound DNA of 80–95 per cent can be achieved after two elutions with 20 µl of distilled water. Stated applications for Isogene include the extraction and purification of DNA from agarose gels, extraction of plasmids from lysates, desalting and/or concentrating DNA solution, separating DNA from enzyme-inhibiting impurities and the replacement of conventional alcohol precipitation or column-based methods for DNA purification. Isogene kits are supplied with DNA binder, NaI solution, washer buffer concentrate — sufficient for 100 assays, says Perkin-Elmer.

Detection systems

Tropix, Inc. has introduced two new optical **luminometers** for use with any chemiluminescent or bioluminescent reaction (*Reader Service No. 111*). The \$4,500 (US) ILA 911 single-well counter is manually operated while the \$12,500 (US) ILA 912 is a fully automated system that can process up to 220 tubes. Both instruments are fitted with a 1.125-inch photon-counting photomultiplier, which, Tropix says, has a 20,000,000 count per second rate maximum and has a programmable optical sensitivity selection. The spectral response range is 320–500 nm. After insertion of the tube in the counting chamber, background measurements, optional injections, shutter aperture selection and integration are automatically sequenced. The built-in microcomputer includes immunoassay, kinetic and DNA probe data reduction protocols and can store up to 30 user-definable protocols. Each luminometer is equipped with a thermal graphics plotter. External computers and data reduction systems can be linked up to the luminometer via the RS232 port.

For the simultaneous **separation of ionic, neutral or polar analytes**, Dionex Ltd

recommends the CES 1 Capillary Electrophoresis System (*Reader Service No. 112*). The CES 1 features a fully programmable dual-polarity power supply with the choice of three control modes — volt, current or power — as well as multistep gradients in all three modes. Also featured are three automated injection modes: electromigration, gravity and pneumatic. A random-access 40-position autosampler provides automatic rinsing of vials and capillaries and buffer changing. Separation takes place in an open capillary using electrophoretic migration and electro-osmotic flow to separate and elute components past an on-column detector. The CES 1 can monitor absorbance at wavelengths between 190 and 800 nm, and



Capillary electrophoresis from Dionex.

a bandwidth of 6 nm ensures detection of samples with injection volumes in the order of 1–10 nanolitres, says Dionex. Programme parameters input via the instrument's keypad can be viewed on the video display and methods can be stored in the non-volatile memory. Hard copy is output in a chromatography-like format.

The collaborative efforts of Intracel Ltd, Nikon and Newcastle Photometric Systems have produced the Epi Fluorescence Systems for **single and dual wavelength fluorochromes** (*Reader Service No. 113*). The Epi Fluorescent System is based on the Diaphot inverted microscope, a series of bolt-on accessories that include photon counters, excitation and emission filter changers, and control software that allows the detection and quantification of Ca^{2+} , H^+ , Mg^{2+} , Cl^- and Na^+ within single cells. For dyes with dual emission characteristics, a beam splitter is used to separate the two wavelengths, and two photomultipliers are used to detect the wavelengths simultaneously. For dyes that require excitation at two wavelengths, filter changers are used in conjunction with a single photomultiplier system. The software provides control of the arc and wheel filter changers and calibration of the photomultipliers. Experimental data can be shown on the computer screen and stored on disc.

Catalogues and services

Over 100 new products are listed in the free 1990 Sigma Cell Culture **catalogue**



Sigma's 1990 cell culture catalogue.

(*Reader Service No. 114*). The 248-page catalogue lists Sigma's complete line of tissue culture products, including 35 new products for plant tissue cultures, liquid and powdered media, Hybri-max products, sera and sera replacements, cell culture-tested reagents, biochemicals and equipment. In addition to the product list, the catalogue includes an expanded section of technical information, including commonly used techniques, stability studies, a guide to antibiotic use, and instructions for supplementing liquid media.

Trouble shoot procedures for transfer methods when using hybridization membranes from Micron Separations, Inc. with their latest **protocol guide** (*Reader Service No. 115*). The guide outlines procedures for Southern, Northern and Western blotting, colony hybridizations, plaque lifts, nucleic acid dot/slot blotting, alkaline blotting, and the use of biotinylated probes with MSI nylon and supported nitrocellulose membranes. The guide features two new sections: how to select an appropriate membrane to suit the transfer procedure and ways to reduce background during the transfer process.

Neuropharm Ltd are offering a **neurotoxicity evaluation service** using the leech brain as a model system (*Reader Service No. 116*). Toxicological screening of compounds has, until now, been based almost exclusively on mammalian studies. Neuro-



Neurotoxicity testing service using the leech brain, from Neuropharm.

pharm are using the simpler nervous system of the leech to carry out acute neurotoxicity tests, which include acute dose response, type of response, mechanism of ion channel and receptor response. Neuropharm, emphasizing the advantages of using the leech, point out that leech neurons are large, easily accessible and individually identifiable. They can be removed as a whole or in part, isolated for several hours and can be impaled with more than one microelectrode.

Custom Immunology is BAbCO's new brochure that describes their **contract antibody development service** (*Reader Service No. 117*). The brochure outlines custom services and protocols provided by BAbCO for polyclonal antibody generation in small and large animals, hybridoma cell-line development, ascites fluid scale-up production of mAbs under GMP guidelines, GLP animal model and veterinary science studies. Other services include antigen preparation, protein isolation and purification, peptide-carrier conjugation, antibody fragmentation, biotinylation, enzyme and fluorescent labelling. □

These notes have been compiled by Diane Gershon, from information provided by the manufacturers. To obtain additional information about these products, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

CUSTOM-MADE & PURIFIED DNA PEPTIDES

\$8/base
50–100 µg
40 hour delivery

Quotes available
50–300 mg
1–3 wks delivery

GUARANTEED QUALITY & QUANTITY

V.G. CORP.

1200 Wonderland Rd. S.
Bldg #7, Unit #2
London, Ontario
Canada N6L 1A8

CALL OR FAX: (519) 652-3758

Reader Service No.29

Custom DNA

Purified and Delivered in 48 hours

100 µg, \$10.00 a base for the first 15 bases, \$7.50 for each additional base. Milligram quantities available.

Research Genetics

2130 S. Memorial Parkway Huntsville, AL 35801

1-800-533-4363

In the United Kingdom 0-800-89-1393

In Canada 1-800-833-8348

FAX 205-536-9016 (NW2971)D

Reader Service No.15

Can student numbers be doubled?

Richard Pearson

The reduced numbers of students in higher education likely to result from demographic changes in the 1990s will have to be countered by changes in attainment and participation rates.

LAST month, the Employment column showed how the output of the UK higher education system has trebled over the past three decades and how this expansion continued even through the previous demographic downturn which started in the late 1960s.

The question for the 1990s is whether the present output of 120,000 graduates a year can be doubled over the next 25 years as suggested by the Secretary of State for Education.

As a starting point, it is important to recognize that while the decade is ending with shortages of graduates, particularly in technological subjects, a significant number of graduates are unemployed and about one in four say they are under-employed.

The demand for good graduates is expected to increase, although it is not possible to forecast precise numbers or the balance between subjects. If demand continues to grow as it has over the past decade, demand could be 30 per cent higher at the end of the century than it is today.

The number of new graduates is expected to increase by more than 5 per cent by 1992. These students are already on their courses. Thereafter the numbers will depend on staying-on rates in education, examination achievements and the attractions of higher education to the young, particularly women, and to vocationally qualified and mature entrants.

Changes

Without changes in these factors, demographic change alone within the social classes will reduce the number of entrants to higher education by 10 per cent over the period 1990–2000.

■ If, however, educational attainment levels rise by a single percentage point, a rate of improvement faster than that achieved in the 1980s, then the number of entrants would be four percentage points higher.

■ If the A-level attainment rate throughout the United Kingdom were to equal that in Scotland, the numbers would be 31 percentage points higher.

■ If the attainment rates of social classes III to V rise to match those of social classes I and II, the numbers would be 65 percentage points higher.

■ If women's participation matches that of men, the numbers should be seven

percentage points higher.

■ If the participation rate of the vocationally qualified rises to match that of A-level students, the numbers would be 11 percentage points higher.

■ Finally, if the participation rate of mature students grows by 50 per cent, the numbers would be 15 percentage points higher in the year 2000.

Scenarios

These changes were combined into two scenarios to illustrate what might be achieved in the next decade¹. The first — rising participation — assumes that attainment rates will rise faster than in the 1980s, participation rates will rise, the number of vocationally qualified entrants will double, the participation rates of mature students will rise by 50 per cent and student loans will not suppress student demand for higher education. Under this scenario, the number of entrants would be 33 percentage points higher in the year 2000.

The alternative scenario, 'constrained expansion', assumes that improvements in attainment rates will continue, but at a rate nearer than of the 1990s, employer competition for school leavers will reduce participation rates, the number of vocationally qualified will be stable and the mature participation rate will rise by 25 per cent. Under this scenario, all the main factors still show improvements on the 1980s, but the effect would be only just enough to counteract the demographic downturn, with student numbers in 2000 being the same as in 1990.

It can be seen that significant changes in attainment and participation rates will be required simply to counteract the demographic effect over the next decade.

To meet the Secretary of State's wish to double student numbers would require changes along the following lines: significant increases in staying-on rates in education at the age of 16 to ensure that A-level attainment rates rise from under 14 per cent to more than 20 per cent, which is almost the level in Scotland; A-level attainment rates of social classes II–V rising by a further half; female participation matching that of males; participation by the vocationally qualified rising significantly to reach the A-level rate; and the mature participation rate increasing by more than 50 per cent.

These represent big improvements on

the rates of change seen in the past decade and would result in fewer young people being available for employment in the labour market.

Competition

Student demand for higher education may, however, be reduced by the growing competition between employers and higher education institutions for the reduced number of school-leavers. Many employers are improving schools liaison and training and career prospects, and are raising salaries in order to maintain school-leaver intakes. They are also widening their recruitment nets to include more women and re-entrants to the labour market. Thus there will be increasing competition to retain pupils in the education system, and to attract both traditional and nontraditional entrants into higher education. The introduction of student loans and a reduction in the size of the maintenance grant is also likely to reduce the short-term attraction of higher education.

The issue for policymakers in higher education, industry and government is how to stimulate student demand and how to organize higher education and provide it with resources to cope with much larger numbers.

Even if student numbers increase, it is likely, assuming that the economy continues to grow, that the demand for good graduates will exceed the supply and that shortages will continue, at least in the first half of the 1990s. Much of this demand will focus on the traditional young graduate with a reasonable class of degree and broad basic skills. Any increase in the supply will include an increasing proportion of nontraditional graduates who, while they may not be 'worse', will be 'different'. A challenge for students, higher education and employers will be to ensure that the attributes of the new supply and the nature of employer demand coincide, otherwise there will be a significant number of graduates who find it difficult to enter what they regard as suitable employment in the 1990s. □

Richard Pearson is at the Institute of Manpower Studies, Mantell Building, University of Sussex, Falmer, Brighton, Sussex BN1 9RF, UK.

1. Pearson, R., Pike, G., Gordon, A. & Weyman, C. *How Many Graduates in the 21st Century? — The Choice is Yours* (IMS, Brighton, 1989).